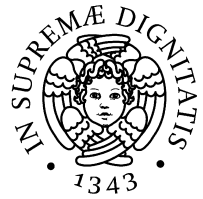

UNIVERSITÀ DI PISA

Dipartimento di Oncologia, dei Trapianti
e delle Nuove Tecnologie in Medicina

Corso di Dottorato in Tecnologie per la Salute:
Valutazione e Gestione delle Innovazioni nel Settore Biomedicale
XXI Ciclo

Ph.D Thesis

IMPROVING DAILY CLINICAL PRACTICE WITH ABDOMINAL PATIENT SPECIFIC 3D MODELS

Vincenzo Ferrari

UNIVERSITÀ DI PISA

Dipartimento di Oncologia, dei Trapianti
e delle Nuove Tecnologie in Medicina

Corso di Dottorato in Tecnologie per la Salute:
Valutazione e Gestione delle Innovazioni nel Settore Biomedicale
XXI Ciclo

Ph.D Thesis

IMPROVING DAILY CLINICAL PRACTICE WITH ABDOMINAL PATIENT SPECIFIC 3D MODELS

Vincenzo Ferrari

Submitted to the University of Pisa in partial fulfillment
of the requirements for the degree of Doctor of Philosophy

Tutor:

Prof. Andrea Pietrabissa:

Dr. Giuseppe Megali:

It was defended on
February 2009

ABSTRACT

This thesis proposes methods and procedures to proficiently introduce patient 3D models in the daily clinical practice for diagnosis and treatment of abdominal diseases. The objective of the work consists in providing and visualizing quantitative geometrical and topological information on the anatomy of interest, and to develop systems that allow to improve radiology and surgery.

The 3D visualization drastically simplifies the interpretation process of medical images and provides benefits both in diagnosing and in surgical planning phases. Further advantages can be introduced registering virtual pre-operative information (3D models) with real intra-operative information (patient and surgical instruments). The surgeon can use mixed-reality systems that allow him/her to see covered structures before reaching them, surgical navigators for see the scene (anatomy and instruments) from different point of view and smart mechatronics devices, which, knowing the anatomy, assist him/her in an active way. All these aspects are useful in terms of safety, efficiency and financial resources for the physicians, for the patient and for the sanitary system too.

The entire process, from volumetric radiological images acquisition up to the use of 3D anatomical models inside the surgical room, has been studied and specific applications have been developed.

A segmentation procedure has been designed taking into account acquisition protocols commonly used in radiological departments, and a software tool, that allows to obtain efficient 3D models, have been implemented and tested.

The alignment problem has been investigated examining the various sources of errors during the image acquisition, in the radiological department, and during to the execution of the intervention. A rigid body registration procedure compatible with the surgical environment has been defined and implemented. The procedure has been integrated in a surgical navigation system and is useful as starting initial registration for more accurate alignment methods based on deformable approaches.

Monoscopic and stereoscopic 3D localization machine vision routines, using the laparoscopic and/or generic cameras images, have been implemented to obtain intra-operative information that can be used to model abdominal deformations. Further, the use of this information for fusion and registration purposes allows to enhance the potentialities of computer assisted surgery. In particular a precise alignment between virtual and real anatomies for mixed-reality purposes, and the development of tracker-free navigation systems, has been obtained elaborating video images and providing an analytical adaptation of the virtual camera to the real camera.

Clinical tests, demonstrating the usability of the proposed solutions, are reported. Test results and appreciation of radiologists and surgeons, to the proposed prototypes, encourage their integration in the daily clinical practice and future developments.

Ai miei nonni

TABLE OF CONTENTS

PREFACE AND ACKNOWLEDGEMENTS

I would like to thank Prof. Mosca for his support to research in computer aided surgery, which allowed me to discover and work in this interesting field. Many thanks to Prof. Pietrabissa and to all surgeons that despite their intensive labour, have found time for answering to my questions and for talking about possible solutions for the improvement of interventions.

I am also thankful to the radiological staff of Cisanello Hospital (Pisa) for their contribution, in particular Carla, which constantly assisted me with her know-how and suggestions.

Many thanks to my colleagues of EndoCAS Center, current and past, that offered me an unequalled mix of competence and friendship. A special thanks to Peppe, my skilful guide during these three years.

Finally I whish to thanks Elisabetta, which trust in my work almost more than me, for her continuous support.

PART 1: INTRODUCTION

1.0 CONTEXT OF THE THESIS

1.1 CURRENT CLINICAL USE OF RADIOLOGICAL IMAGES

Hospital-based care is a complex therapeutic environment with many different compartmental and interconnected activities of the various departments with its own protocols, teams and goals. All the major published data on hospital-based morbidity indicate in communication between the various department and in the interpretation of device data information one of the prime setting for these mishaps [1] [2]. In radiology, recent developments in CT and MRI equipment and the introduction of new contrast medium allow the acquisition of volumetric datasets, relatives to human anatomy, functionality and pathology, with high degree of detail. This amount of information is difficult to interpret because generally an exam is composed by many phases, each with hundreds of images, and its interpretation requires the knowledge of how the contrast media flows inside the anatomical structures and the 3D interpretation of the anatomies described by the volumetric dataset. Radiologists build a "mental model" of the specific anatomy using their anatomical knowledge and examining in general slice by slice the volumetric dataset. Furthermore the detailed information contained in a volumetric dataset are partially lost passing from the radiological department to the surgical room. The surgeon usually plans the intervention just interpreting the information provided by the radiologist, consisting in the diagnosis coupled with only few 2D images selected from the volumetric dataset.

1.2 FROM RADIOLOGICAL IMAGES TO PATIENT SPECIFIC 3D MODELS

The new surgical scenario, based on the "computer assisted" model, allows reducing these troubles. There are possibilities and requests from physicians for the visualization of 3D datasets with virtual models that would simplify the reading of exams and remove errors from their interpretation, offering great benefits in the phase of diagnosis and surgical planning. Moreover surgical simulators, that today allow the surgeon to perform routinely tasks just on a standard anatomy, could be improved with the possibility to work with the

specific model of each patient. In this way the surgeon could try the more difficult tasks of each intervention in a virtual environment that simulates the real surgical scenario, choosing the best surgical strategy in advance. Furthermore, the integration of anatomical virtual models with traditionally and robotic/mechatronic devices in the surgical room will result in fundamental changes in therapeutic strategies and approaches. In fact image guided surgery, robotics, bioengineering, augmented-reality, and intelligent surgical tools are contributing to evolve the surgical procedures from a "traditional" model, based only on the surgeons' skill, to the "computer assisted" model, where surgeons and machines work together in a synergistic way. The surgeon can use systems that permits him to see covered structures before reaching them, he can see the surgical scene (anatomy and instruments) from different point of view using surgical navigators, he can use intelligent instruments that assist him removing tremor, correcting his movements or doing repetitive tasks like ligatures[3] [4]. All these aspects will be useful in terms of safety, efficiency and financial resources for the physicians, for the patient and the sanitary system too.

There are many acronyms that include, totally or partially, the previous concepts: CIST (Computer Integrated Surgery and Therapy), CAMI (Computer Assisted Medical Interventions), IGS (Image Guided Surgery), IGRT (Image Guided Radio Therapy), CAD (Computer Aided Diagnosis), CAS (Computer Assisted/Aided Surgery), and many others. In this work it is used CAS when the system or the concept is mainly addressed for surgical assistance, while it is used CAD when the focus is the diagnosis.

The dissertation is mainly focused on the generation and the use of patient specific 3D models in the clinical practice, especially regarding mini-invasive abdominal interventions, but deals also issues that could be useful for CAS/CAD applications for other anatomical districts.

In the next paragraphs are shown some examples of potential applications using patient specific 3D models that are addressed in this thesis work: 3D anatomy visualization, surgical navigation, smart mechatronics surgical tools and mixed-reality. However the issues treated in this work could be useful for others types of applications. For example, the development of biomechanics or physiologic models, that requires the geometric knowledge of the organ to simulate and its positioning inside the surgical room.

1.2.1 3D anatomy visualization

3D visualization offers benefits for all clinical departments that are involved with radiological images visualization. In order to appreciate the potentialities of the 3D visualization consider Fig. 1.

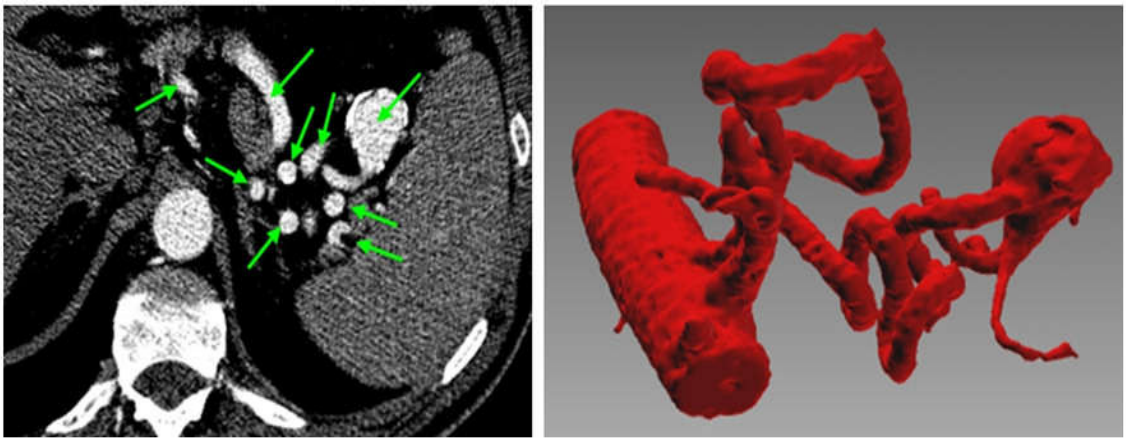

Fig. 1 (left) Traditionally 2D visualization in a CT slice of a splenic artery indicated with the arrows. (right) Corresponding 3D model.

In case similar to this one, also a radiologist, watching only the 2D images slice by slice, has difficulties in the interpretation of the 3D representation of the anatomy. Furthermore this type of visualization allows to view and interact with complete models, like the one in Fig. 2. Using these models it is possible to rotate and zoom the scene, to turn on/off the image of the various organs or to change transparencies, to response to light producing shadows, ect. It allows to subjects, not familiar with the visualization of medical imaging, to understand the topological and geometric significance, and simplifies the interpretation work to radiologists, and in general, to physicians.

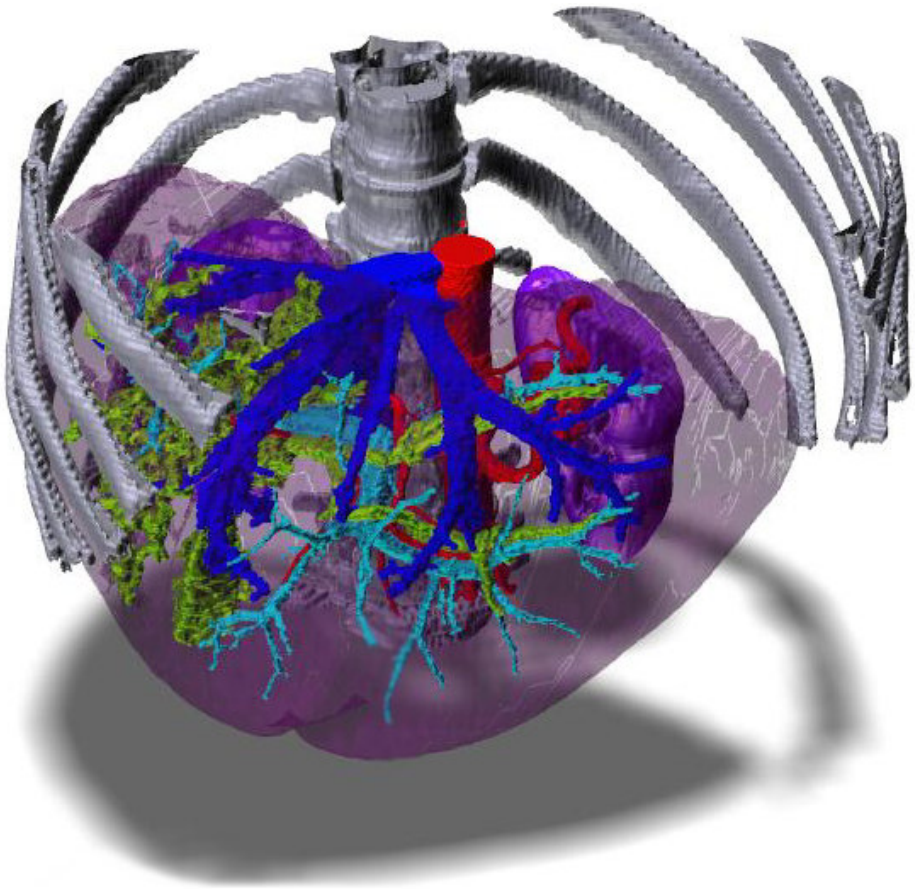

Fig. 2 Patient specific 3D model generated from a CT dataset.

1.2.2 Surgical navigation and smart mechatronics surgical tools

Other typologies of CAS and CAD systems can be built aligning the pre-operative information relative to a patient (obtained by means of radiological devices) to the intra-operative information in the surgical room, consisting in the real patient and traditionally and supplementary devices.

In the next figure is shown the potentialities of surgical navigation, where the surgeon can see a virtual scene, with virtual surgical instruments and virtual patient aligned with the real instruments and patient.

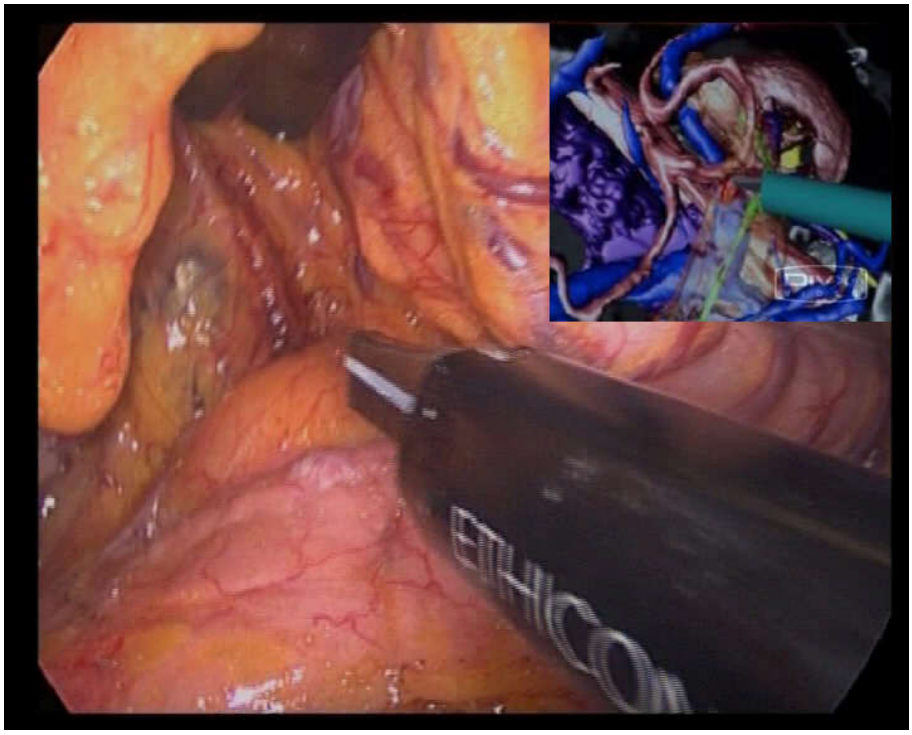

Fig. 3 EndoCAS Laparoscopic Navigator at work. The laparoscopic surgeon work using the traditional real endoscopic images. In addiction he/she can see, on an additional monitor the virtual scene (up-right), with the surgical instruments that are moving respect to the virtual patient.

Surgical navigation enhances surgeon's perception:

- ✓ It allows to understand the relation between the instruments and covered anatomical structures.
- ✓ It allows to view the scene from different point of view.
- ✓ It allow to remove all parts not important for the current task from the virtual scene enhancing the understanding.

Using pre-operative information as guide allows also the development of smart mechatronis systems that, "knowing" the anatomy, can assist actively the surgeon during the intervention. The image-guided concept is adopted in IGRT with the re-positioning of the patient [5] respect to the planning image, for the positioning of needles to perform biopsies, like in the example in Fig. 4 and in [6] [7], in orthopaedics and in dental implantology.

Fig. 4 Experimental setup for robotic image guided biopsy at EndoCAS.

Current robots used in laparoscopic (e.g. the da Vinci Surgical system) are master-slave teleoperation devices with no automation besides tremor filtering. These robots, jointly to the use of patient specific 3D models, could assist the surgeon with advanced functionalities. In an early future, functionality like reducing instruments velocity, when the surgeon is working close to vital parts, up to (eventually) block the instruments, to avoid dangerous contacts, could became reality.

1.2.3 Mixed-reality

The last example of functionality regards the mixing of the real and the virtual scene obtaining a mixed-reality. An example of this concept is shown in the next figure.

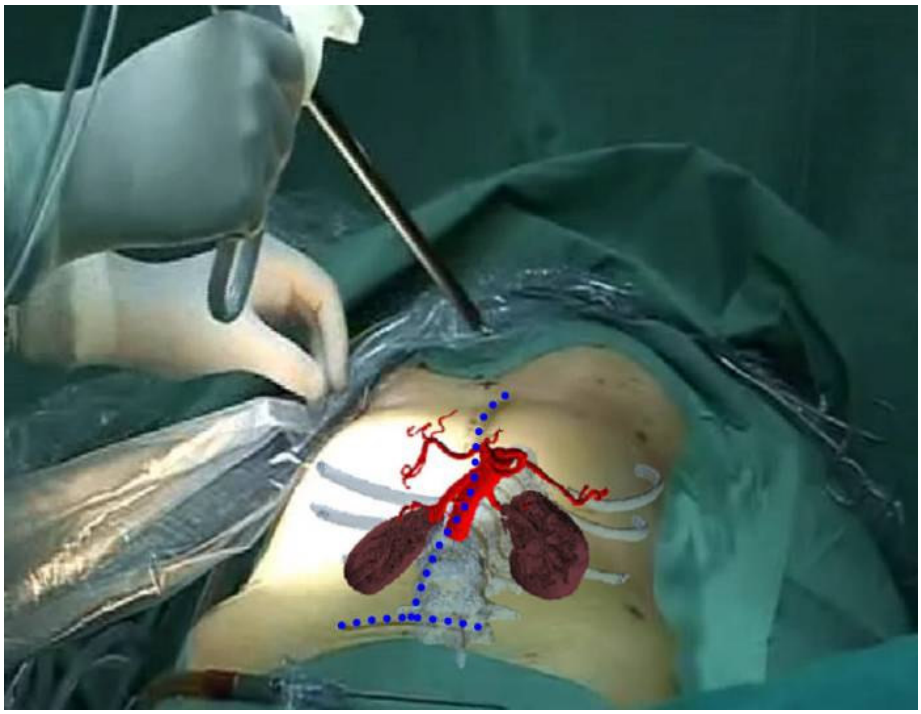

Fig. 5 Mixed-reality in the operative room.

This functionality introduces many advantages for each task where the physician have to interact with the patient (palpation, introduction of biopsy needle, catheterization, intervention, etc.).

Similar concepts are called augmented-reality, when a real scene is added by some details given from the virtual information, and augmented-virtuality, when the most important information is virtual and only a minor part is real. In our case virtual and real information can be considered with the same degree of importance, in terms of physics dimension and in terms of value for physician work, so it is better to talk in general terms of mixed-reality [8].

1.3 ADVANTAGES

How shown and partially explained in the previous paragraphs, the CAD and CAS concept introduce many advantages for the clinician. Today computer aided solutions have been developed and used for specific diseases in few interventions. Their use for the abdominal district is limited to virtual colonoscopy, some types of radiostherapy and few others examples. The advantages that a massive use of patient specific 3D models would introduce for the diagnosis and the treatment of abdominal diseases are shown below.

Radiologists could benefit of patient anatomy 3D visualization, obtaining a simplification for the interpretation and reducing the diagnostic errors. The same models could increase the communication between the radiologist and the surgeon, between the radiologist and the general doctor and also between the physician and the patient. In general 3D visualization allows the exchanging of medical information between figures with different anatomical and radiological knowledge degree. Furthermore 3D models introduce new fine training aids for the study of the anatomy and for a better understanding of diagnostic and therapeutic strategies. Regarding training, a new generation of patient specific surgical simulators could support the young and also expert surgeon. The young surgeon could try many surgical tasks on various real anatomies, with clear advantages respect to work with a fixed standard model like in current simulators. The expert surgeon, how introduced in 1.2, could try in advance the simulation for a specific real intervention. The development of patient specific surgical simulators requires, over than the generation of the specific 3D models, additional work from computer scientists. However they could be real in few years. While, 3D visualization of the patient's anatomy, in the planning phase and during the intervention, could offer to the surgeon a better understanding of the anatomical working structure, just tomorrow. Surgical navigation, robotics and mixed-reality could also decrease the time required for interventions and surgical errors. The invasiveness of the intervention could be reduced too. In fact, the definition of the optimal access port and the visualization of covered structures could limit the damages to healthy tissues. Finally, 3D patient models could be useful for the entire patient clinical workflow, up to control visits after the intervention. When the physician have to verify the status of a known disease inside the patient body, the first step is to localize the position of the disease, that could be facilitated using virtual navigation aids. An example of a system for this purpose using US can be found in [9].

All the advantages offered by CAS and CAD technologies leads to a significant improvement of quality in surgical/healthcare services contributing to the social welfare and safeguarding citizens' wealth. The most important aspects for the patient, for the doctor and for the sanitary system are:

- the reduction of the intervention invasiveness allows the patient to return quickly to his/her normal life;
- the reduction of morbidity and mortality thanks to the decrease of faults caused by erroneous cutting of indistinguishable vital anatomical elements. This aspect is particularly important in minimally invasive surgery (MIS) where the surgeon operates under perceptual and motor constraints due to visual and tactile limitations [10] [11];
- the improvement of the precision of the diagnosis that allows to reduce costs due to ineffective therapies and costs for legal causes required from damaged patients.

I reported advantages of the CAS and CAD paradigm only in general terms. Specific advantages depend on the particular application. I noted that potential applications, with the correlated advantages, can be found in each clinical sector where the diagnostic or surgical procedures are difficult and depends on the geometry and on the topology of the anatomy. New useful applications can be found talking with doctors that work every day for a specific problem. Some times the solution is proposed by the CAS/CAD expert, while in other cases it is directly the doctor that find the potentiality of the computer aided approach for his/her work. In each cases, after the definition of the new approach for a specific problem, the doctor endorse the development of the just defined CAS/CAD system, which in some cases become real.

1.4 TROUBLES FOR CAS/CAD SYSTEMS INTRODUCTION IN THE DAILY CLINICAL PRACTICE

Even though the new computer aided clinical scenario is explored by more than one decade and many prototypes have been built, a massive employ of CAS and CAD technologies is quite far from to be real and there are some troubles that have to be overcome for their introduction in the daily clinical practice. There are shown now the main concepts for

building CAS/CAD systems [12], coupled with some consideration regarding their use in the abdominal district and their integration in the daily clinical practice.

1.4.1 3D models generation

Commercial radiological software suites offer in general 3D reconstruction tools. For example maximum intensity projections (MIP) is commonly used in diagnosis. It consists in the 2D projection of a portion of the acquired volume and allow, positioning the projection plane and the volume limits, to represent in a single 2D image many 3D information. It is useful especially for representing the vascularization in a limited volume. In case of large volumes the information will be lost, due to the superimposition of the various structures. For those cases, the radiologist can visualize the scene using direct 3D volume rendering tools. They allow to rotate and zoom the virtual scene allowing the user to work with a realistic and natural visualization support (Fig. 6).

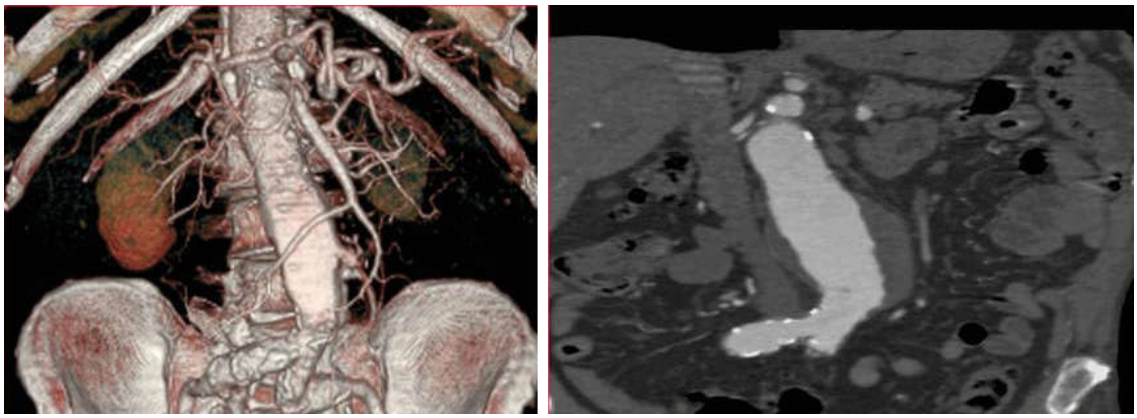

Fig. 6 Volume rendering example (left) from a CT dataset (right).

This type of visualization is very useful and simple to use. It requires mapping each intensity value of the dataset to a colour and to an opacity values. The user can choose standard transfer functions, or he/her can define a particular one using ramps, piecewise linear functions or arbitrary tables. Standard transfer function windows are pre-designed for the visualization of the various anatomical structures. However for some particular cases, direct volume rendering do not allow to represent correctly the 3D information. For example the volume rendering image of Fig. 6(left) do not show the aneurysm around the abdominal artery visible in Fig. 6 (right), because the blood near the wall of the vessel is coagulated,

consequently it is not contrasted, and so it is not visible using the standard transfer function for arteries. Unfortunately trying to include in the window the values of the coagulated blood, many others voxels, with the same intensity, will be rendered and the scene will be indistinguishable. Furthermore, since direct volume rendering is based on the intensity and not on the nature of the voxels, they do not allow to turn on/off models of the various organs in the scene or to change their transparencies, their response to light, in order to produce verisimilar shadows and reflects, and finally they do not allow the simulation of the interaction of real or virtual surgical instruments with the anatomy. All these functionalities require the segmentation of the dataset for the morphological characterization of each anatomical structure, which, thanks to the established algorithms based on marching cubes [13], allows 3D surface model building (Fig. 7). This type of visualization is also faster than direct volume rendering, so real time rendering is possible also with large volumes.

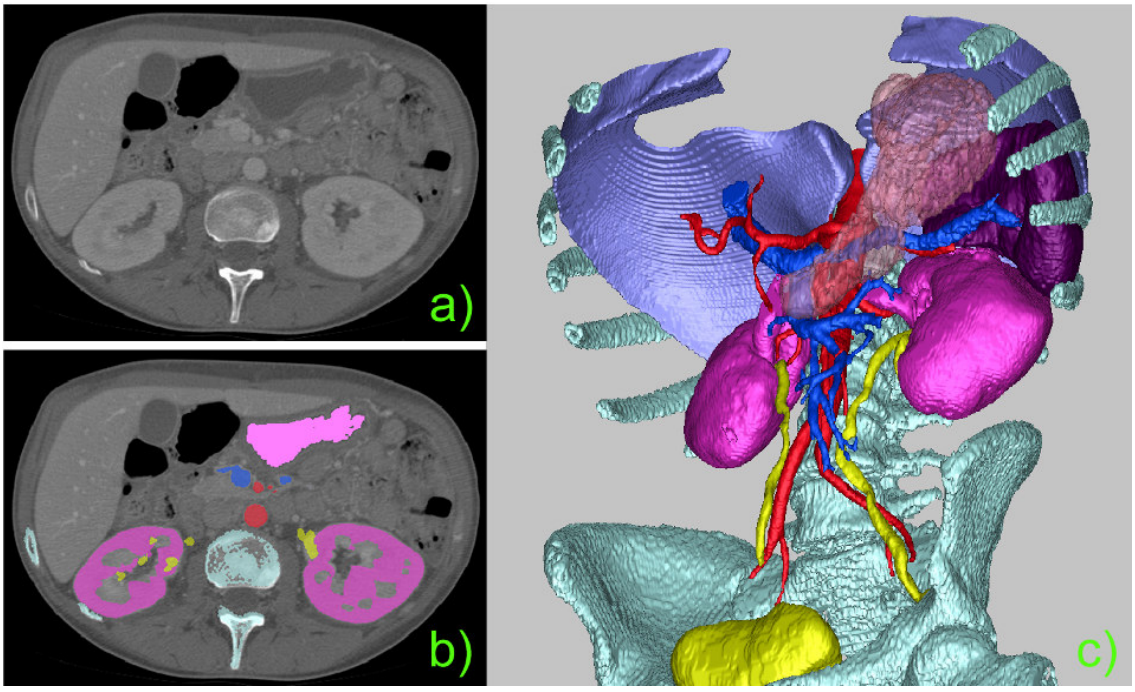

Fig. 7 Modeling of anatomical structures: (a) CT dataset slice, (b) Segmentation and Labeling, (c) 3D surface models extracted from segmented and labeled images.

Automated and accurate segmentation is one of the "holy grail" of CAD and CAS [14]. Even if there are many works in the literature for general purpose segmentation [15] [16], real applications are limited for some specific anatomical structures.

Some companies and research institutes propose automatic segmentation software products that provide good results on non abdominal anatomical structures. Regarding abdomen, some new research works propose good results for some lower abdominal structures [13] and some promising works [14] [15] propose to segment all the others. But unfortunately, today, automatic segmentation tools for the abdominal cavity do not guarantee enough results for each organ in terms of precision and accuracy.

Semiautomatic segmentation tools, usable in the clinical practice, can be found for the liver [17] [18] and for the colon (virtual colonoscopy).

The segmentation solution for the other abdominal organs, that guarantee enough accuracy, is the application of general purpose suites based on a collection of segmentation tools that the user can use freely. These suites are difficult to use from medical users because they require an optimal mathematic knowledge of the various algorithms and experience in their usage. In fact, these suites require that the user set a lot of mathematical parameters for each method and define the right sequence of methods to apply on the original image data. For this reason their use requires that the radiologist, who "knows what to do" and is able to validate the result, is assisted by a technician, who "knows how to do" and is able to use that kind of software. This solution is not applicable in the clinical scenario for economics reasons. It requires the work of two professionists for some days.

1.4.2 Localization

Localization allows real-time tracking of position and orientation of moving (or fixed) objects (such as surgical tools, robotic arms, and patient), permitting to realize interactive navigation systems that augments the information for the surgeon, using virtual instruments and mixed-reality views, and enhances his/her performance, using smart mechatronics surgical tools, during the intervention. Various localizers, with optic or magnetic technology (principally), are available in the market.

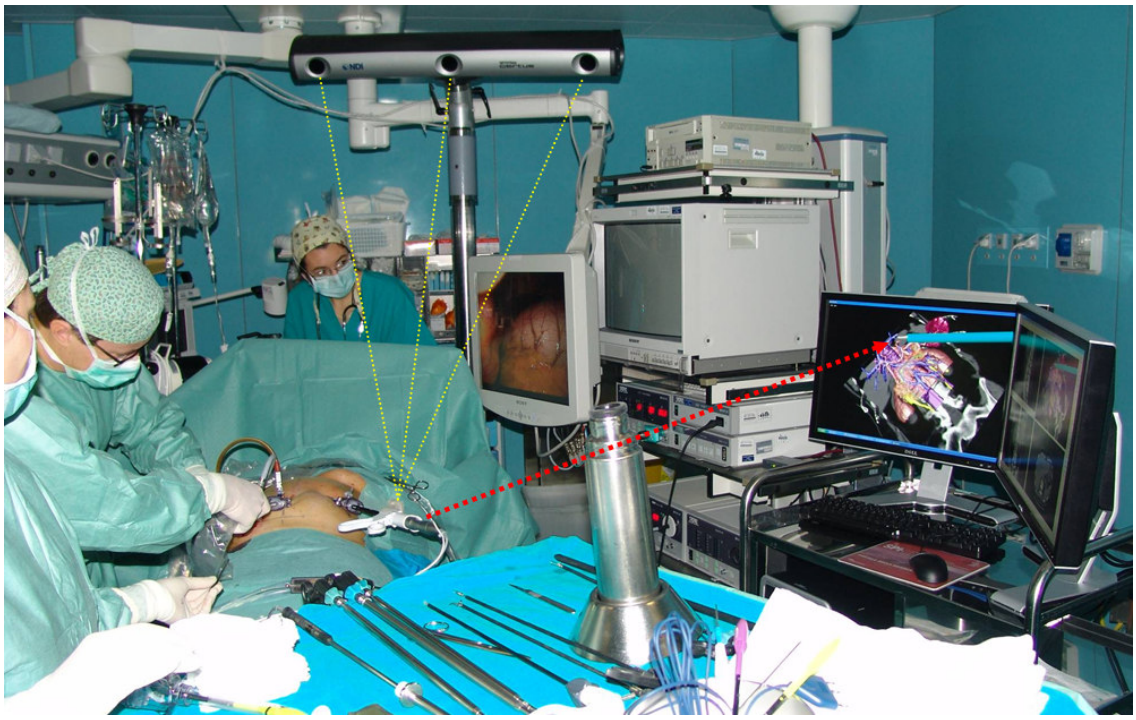

Fig. 8 EndoCAS Navigator in the surgical room. Surgical instruments are tracked by an optical localizer.

Localization of traditionally tools and instruments requires their sensorization and calibration. Sensors have to be designed and positioned in order to guaranty their functionality and safety (expecially regarding the sterilization). The calibration procedure has to be done to determine the relation between the sensor and the tool silhouette. These two tasks can be difficult for prototype developing (see for example US calibration [19]), but would be very simple if sensors are integrated in the tools.

In general commercial localizers require frequently calibrations, can have a large footprint and can be expensive (up to 80k €). However they offer the localization accuracy necessary for many CAS and CAD applications. Their integration in the surgical rooms has to be taken into account, at least by reference hospitals, otherwise the introduction and the testing of new computer aided systems will be drastically limited.

On the other hand we can think to applications that not require the use of an external localization system. Some times the same problem can be solved with or without the use of an external tracker. For example, in detal inplantology, some companies localize the drill and the patient for the assistance in drill positioning (respect to the planned trajectory on the CT images). Others company propose custom-made stereolithographic surgical guides, based on the simulated implant, that provide highly accurate drilling template that fits perfectly on the

mouth and ensures safe and predictability, so avoiding the needs of trackers. The second solution is the most diffused for its simplicity and the minor cost.

1.4.3 Patient registration

Mixing virtual pre-operative information (extracted from the medical dataset) with real intra-operative information (consisting in the patient himself), requires the registration step: the alignment of the virtual anatomy to the real one. This task requires to determine the geometrical transformation of correspondent points taken in two different reference frames and in two different time instants. In fact, pre-operative information are given in the reference frame of the radiological device and are acquired some days before the intervention, while the intra-operative information are given in the reference frame of surgical room (defined by means of a tracking system) and are acquired during the intervention.
Many works have been proposed [20], but for movements of the abdominal and torso structures, real CAS applications are limited to bones (ortopedics) and head (neuro surgery).
Registration could be simplified and eventually skipped using volumetric intra-operative imaging, but unfortunately the surgical room has in general no radiological volumetric devices.

1.4.4 Mixed-reality implementation

The fusion of real world elements (grabbed by means of cameras) with virtual synthetic elements in a single view requires, in order to mix coherently the real with the virtual information, at first the localization of image source and real objects with the correlated problematic described in 1.4.2. Further, the projection model of the virtual camera (in general implemented in a scene graph library), has to be adapt to the real camera. These problems can be solved today using image vision routines, but requires particular attentions regarding their use with traditionally surgical cameras like laparoscopes. Often it is necessary that the surgeon do not change the camera configuration, for example in term of zoom or focus, during the intervention. These limitations, coupled with the depth perception limitations using single camera view, reduce the applications of mixed-reality in the clinical practice.

2.0 CONTRIBUTION OF THE THESIS

2.1.1 Work approach and the EndoCAS Center

The work was done at EndoCAS center in the Cisanello hospital of Pisa (Italy). One of the main activity of the center is the development of high-tech systems designed to overcome the current limits of surgery and radiology. EndoCAS carry out simultaneously basic and applied research. Starting from real clinical problems and defining the technical-functional specifications for an "ideal" system that can solve them, the center faces the basic research issues to find the solution necessary to develop the system. In the opposite direction, the results of basic research at the state of the art are pushed into the design of new CAS systems in order to improve the current surgical procedures, to reduce their invasiveness, or to allow new interventional procedures.

Following this approach, the thesis work deals the treatment of abdominal organs, using patient specific 3D models, taking deeply into account the usability of the solutions in the daily clinical practice and using an interdisciplinary working method. The solutions proposed take attention to the specific problem to solve and to the anatomical structures involved considering all the available technologies (looking ahead to the early future), and the possibility to integrate them in the current patient workflow in order to find the more simple solution to the problem.

2.1.2 EndoCAS Navigator

Many of the proposed solutions were integrated and used in the EndoCAS Navigator platform. In other cases EndoCAS Navigator was used as testing environment because, integrating several aspects of CAD and CAS into a modular open architecture, allows rapid developing of new functionalities. The dissertation often refers to EndoCAS Navigator platform and its components.

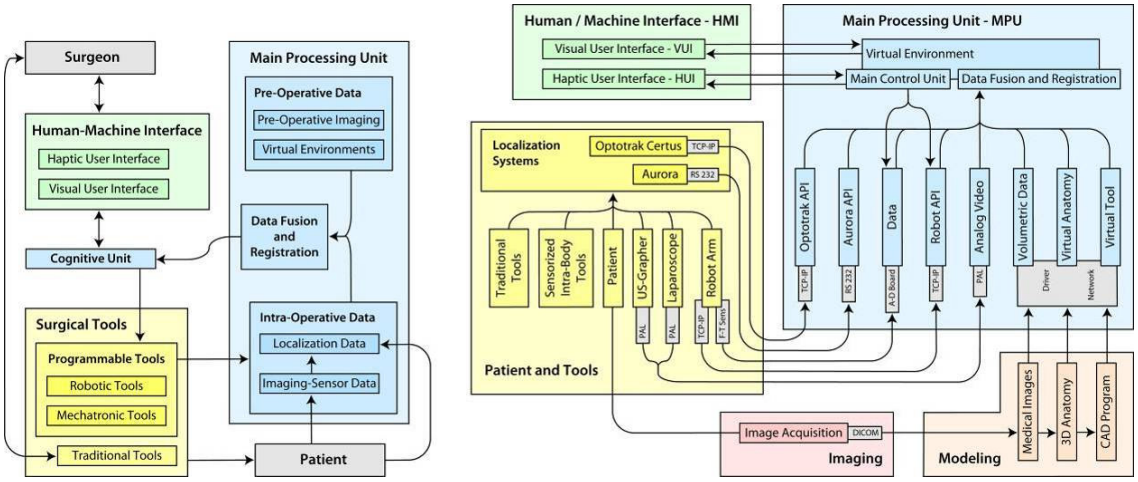

Fig. 9 Functional scheme of the computer assistance system (left) and scheme of the integrated EndoCAS Navigator system showing: the hardware and software components, the architecture and the intercommunication (right).

From a functional point of view, the specifics of the platform are illustrated in Fig. 9 (left). The scheme highlights the communication between the main functional modules of the system and the interaction between system, surgeon and patient. The platform consists of three main functional modules: the surgical tools, the main processing unit, and the human/machine interface. The surgical tools module comprises the instruments used to perform the interventions. Tools are classified into traditional tools and programmable tools. Tools commonly used in surgical practice and managed by surgeon in a traditional way fall in the first category. These tools, used for imaging (laparoscopes, ultrasound probes, etc.) and intervention (scalpel, forceps, cauterizer, drill, biopsy needle, etc.), are passive, for what concerns movement control, and work under direct manual control of the surgeon. In contrast, programmable tools category encompasses active, intelligent tools (such as mechatronic and robotic tools), provided with sensors and programmable actuation capabilities.

The main processing unit (MPU) processes and integrates preoperative data with intra-operative data concerning the surgical environment and the internal status of the programmable tools. Integrated data (provided by the Data Fusion and Registration module) are processed by the Cognitive Unit and returned to the surgeon in form of sensorial enhancement by means of the Human/Machine Interface (HMI). The HMI is composed by two modules that can function independently: the Visual User Interface (VUI) and the Haptic User Interface (HUI). The status of both interfaces is updated in real-time.

The surgeon interacts with the programmable tools through the HMI. The Cognitive Unit, integrating commands given on the HMI with the information provided by the MPU, provides for visual safe guidance and monitoring dangerous situations that may occur during navigation (i.e.contact, proximity etc.) and acts as an intelligent and active filter to the programmable tools commands given by the surgeon, inhibiting or re-interpreting the most critical ones. The synergy between system and surgeon is achieved by means of the Cognitive Unit which by implementing a closed loop between surgeon's commands, programmable tools and MPU, enhances overall performance.

EndoCAS Navigator is based on the described functional approach, and enables the selection of the appropriate components for specific applications. The system can be used for pre-operative visualization, diagnosis and planning, intra-operative passive and active guidance. Furthermore, the system integrates components such that it is capable of adaptation for a variety of application domains. The integrated system is illustrated in Fig. 9 (right), which highlights the hardware and software components and their intercommunication.

The availability of virtual models of all relevant elements in the surgical scene is a pre-requisite for the construction of the Virtual Environment. Medical images of the patient are acquired preoperatively (Image Acquisition). Surface models are created by a modelling process (described in chapter 3 of this thesis) in order to build realistic geometrical virtual models of the anatomical organs and structures (Virtual Anatomy) involved in the intended operation. Virtual models of the surgical tools (Virtual Tools) and of all devices that will interact with the patient are generated using computer aided design programs.

During the intervention, in order to place the elements correctly in the surgical scene, real-time information about their spatial position are provided by the localizer. The different reference frames, in which spatial coordinates are described, need to be co-registered and aligned with the virtual representations of the anatomies (registration, described in chapter 4).The geometrical description of the surgical scene is enhanced by information derived from intraoperative imaging devices (Laparoscope, US) and data collected by different types of sensors. All these data sets are integrated into the virtual environment by a Data Fusion process. Both optical (Optotrak Certus®, Northern Digital Inc.) and electromagnetic (NDI Aurora®, Northern Digital Inc.) localization devices have been integrated in the platform respectively for external-body and internal-body localization. A software module, on the top of API of the localizers, that provides a unique interface to configuration and management functions, and allows the use of both in the same application, has been developed and integrated. The module also implements methods for calibration of localization sensors with

respect to tools shape and functionalities. Specific procedures have been implemented for automatic dynamic calibration of sensors mounted on the surgical tools, and for manual calibration based on the digitalization of reference points on the tools. Other calibration procedures concern the robot-localizer calibration, and intra-operative imaging devices calibration (such us laparoscopic camera and US probe). The control loop implemented in the core of the MPU (Cognitive Unit) monitors the virtual environment and is responsible for determining the feedback actions associated to the state of the virtual environment.

Virtual environments are created integrating in the same view both extracted surfaces and original volumetric datasets (orthogonal slices). The visualization module (developed using the open source framework OpenSG [25]) allows the visualization of virtual environments, modification of the virtual scene settings (transparency, slice position, organs to be visualized), virtual navigation inside the patient by moving the viewpoint by means of a 6D mouse, and perception of stereoscopic images by means of a Head Mounted Display (HMD). Also mixed-reality functionalities have been integrated. The module implements two main functions: the video acquisition and streaming function that manages the image capture from a generic local or remote video source, and the mixing function that synthesizes the hybrid image using video frames and virtual 3D models (chapter 5 of this thesis).

In the platform has been integrated an industrial robot (Samsung FARA AT2) to provide active surgical assistance and accurate positioning during intervention [27]. The robot reference system has been calibrated with the global reference system (given by the optical localization system) so that the robot can be moved along planned trajectories in closed loop with the surgical navigator. An automatic iterative calibration method has been implemented. It is based on the Lie algebra [28], to be performed at the beginning of surgical intervention and every time the relative position between the robot base and the localization system changes. Since the usage of the robotic arm is limited to only few steps of the surgical intervention, a mechanism, based on a 6DoF force sensor (Mini45, ATI Industrial Automation, Inc.) mounted on the end-effector of the robotic arm and an admittance controller allow the surgeon to move the robot (in and out from the surgical scenario, or to a precise position) just by exerting force at its distal part.

2.1.3 Roadmap

The **chapter 3** of the work gives a method for the 3D segmentation of some abdominal structures, for which exhaustive studies have not been yet proposed. The obtained procedure was designed and tested on 20 cases given by standard CT exams and it is focused on abdominal aorta, vena cava, portal vein, splenic vein, splenic aorta, ureters, kidneys, pancreas, spleen and bones. The design process of the segmentation tool was guided from the objective of building a fast sequence of operations, in terms of manual operations and in terms of computational time, in order to reduce the total segmentation time and to allow its use by radiologists during the diagnosis work.

The **chapter 4** regards the registration problem of the abdominal cavity for CAS interventions. The work examines the various categories of errors through anatomical considerations and results of experiments made in my laboratory and in others works. The experiments were made studying the motion during breathing of the abdominal wall and the inaccuracies introduced by different decubit, using an optical localizer, and the motion of internal organs, using CT, MRI and US. It is reported some suggestions for the preliminary rigid body registration to use directly in the surgical room or to use as starting point for more accurate deformable approaches.

The work descript in the **chapter 5.0** starts from the consideration that working without the use of volumetric intra-operative imaging, we encounter a lack of information for the modelling of all abdominal deformations, necessary for a complete non rigid registration task. A large amount of intra-operative information could be acquired with the elaboration of laparoroscopic images. Therefore they were evaluated monoscopic and stereoscopic 3D information extraction approaches, using laparoscope and cameras in general. It allowed also to introduce other potentialities for the CAS scenario. First, regarding mixed-reality, the projection model of the virtual camera has been analytically adapted to the model of the real camera, obtaining a perfect alignment of the virtual view to the real view. Second, they have been developed solutions that do not require the use of an external localization system, where camera position is auto-localized in the scene elaborating the images.

PART 2: WORK DESCRIPTION

3.0 ABDOMINAL CAVITY SEGMENTATION

This chapter descripts a segmentation approach for MDCT (multi detector computed tomography) post-contrastographic images on the base of anatomical, empirical and functional consideration concerning the timing of contrast distribution. It is based on the use of the optimal phase for the segmentation of each abdominal structure and on an optimal removal sequence that, starting with easy structures, allows to identify objects hard to segment. Radiologists, using a software tool ad-hoc developed, based on region growing, following the proposed extraction sequence can generate and validated 3D models for bones, ureters, kidneys, pancreas, spleen, arterial, portal-splenic-mesenteric and cava vessels in about 30 minutes of work.

This part of the thesis has been done with the collaboration of the radiologist Carla Cappelli, a PhD student of the course "Tecnologie della salute: valutazione e gestione delle innovazioni nel settore biomedicale" offered by University of Pisa, as me.
The work is supported by the "Cassa di Risparmio di Pisa" Foundation.

3.1 INTRODUCTION

Segmentation process discriminates the anatomical structures (organs, vessels, etc...) and assigns each voxel of the radiological volumetric image to the relative virtual object (the assigning process would be called "labelling", but today the term "segmentation" includes implicitly this process too). Knowing voxel dimensions it is consequently possible to describe geometrically each segmented region.
The service offered by automatic and semi-automatic segmentation algorithms are analogous to the mental work made by radiologists reading a dataset. In fact the radiologist has to recognize voxels, related to the various anatomical structures, often watching slice by slide 2D images, visualizing the relative 3D representation only in his mind.
Like described in the introduction the segmentation of a dataset introduces many useful functionalities. Thanks to the established algorithms based on marching cubes [13], the 3D

surface model of the organs of a patient can be easy built, starting from a segmented dataset, obtaining the advantage of a fast rendering, especially on traditionally graphics card that are optimized for working with polygonal mesh surfaces. Furthermore, since these models are based on the nature of each voxel (and not on the intensity), they allow to turn on/off models of the various organs in the scene, or to change their transparencies, their response to light (in order to produce verisimilar shadows and reflects), and they allow to recognize the interaction of surgical instruments with the anatomy. Finally, segmentation, characterizing morphologically each structure, allows developing patient specific surgical simulators, which the surgeon can try preventively the various interventional techniques, and allows their quickly physics realization using rapid prototyping technologies (useful especially in orthopaedics, dental implantology and plastic surgery).

3.2 STATE OF THE ART

Even if there are many works in the literature for general purpose segmentation [15] [16], real applicable solutions for the abdominal district are limited.

In fact some companies and research institutes propose automatic segmentation software products that provide good results only on specific, non abdominal, anatomical structures. In general algorithms used by them are based on basic threshold or region growing techniques [21] [22], active contours [23] [24] and anatomical atlas [25]. For example Materialise (www.materialise.com) offers good results with Mimics on bone structures using basic techniques, while Cerefy (www.cerefy.com) offers some products for cerebral images using its brain atlas. Regarding abdomen, some new research works propose good results for some lower abdomen structures [26] using deformable models, while some promising works [27] [28] propose to segment abdominal organs using probabilistic atlas. But unfortunately, for the abdominal cavity, automatic segmentation algorithms do not guarantee enough results for each organ in terms of precision and in terms of accuracy.

About semi-automatic software, there are some works for the segmentation of the liver and for the colon (for virtual colonoscopy). For example the system developed at ICG (Graz University) [17] allows the radiologist to refine the liver parenchyma using augmented-reality tools after a rough automatic extraction. Mevis (www.mevis.de) offers a service for

hepatic surgery too, where a team of radiologists uses semi-automatic algorithms [29] for building 3D models from datasets sent by customers.

Until now, the segmentation solution for the others abdominal organs, applicable in the clinical scenario, was been to apply general purpose suites based on a collection of segmentation tools that the user can use freely (like Analyze by Mayo Research, Amira by Mercury Computer Systems or Osirys by University Hospital of Geneva). These suites are difficult to use from medical users, because they require an optimal mathematic knowledge of the various algorithms and experience in their usage. In fact, these suites need that the user set a lot of mathematical parameters for each method and define the right sequence of methods to apply on the original image data. For this reason their use requires that the radiologist, who "knows what to do" and is able to validate results, is assisted by a technician, who "knows how to do" and is able to use that kind of software.

3.3 APPROACH

Modern MDCT scanners allow obtaining quickly accurate volumetric images of the patient in several distinct phases after the injection of the contrast medium (arterial, venous, delayed phases, etc). The use of multiple phases is useful because different anatomical structures enhance in different and typical times after the injection of the contrast medium (for example arteries enhance about 25 seconds after the injection). Consequently radiologists, applying consolidated protocols, employ frequently multi-phase MDCT datasets, because they offer additional information respect to un-enhanced (basal) images.

Formalizing actions and mental processes done by radiologists, during the observation and analysis of medical images, it has been designed and implemented an efficient segmentation procedure that allows the creation of 3D models of abdominal anatomies in a simple and rapid way.

The procedure uses more than one post contrastographic phases, as radiologists in their work, because it is simpler to extract each organ from the phase where they are better visible. Another idea, derived from radiologist mental reconstruction, is to treat initially structures that are easy to segment (i.e. clearly identifiable on the images), and to remove them from the dataset. This approach is similar to that one used by each of us when we discover a new

environment in our life. First we search reference points basing on our experience and then, basing on these known references, we try to discover new objects, paths, levels and so on. Using this method, the treatment of structures difficult to segment is simplified like proved in [28] [30]. An example of the benefits offered from this technique based on the extraction of pre segmented structures is reported in Fig. 10 .

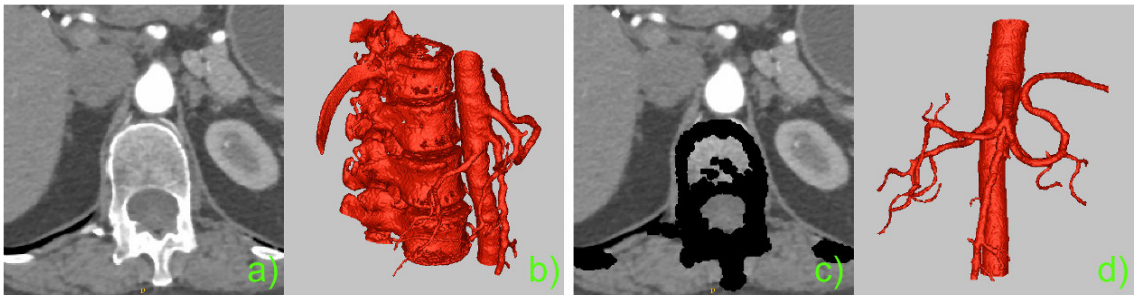

Fig. 10 (a) CT slice of an arterial phase: lumbar arteries touching the spine and having the same density could compromise the segmentation, in fact the 3D model of the artery obtained directly from the dataset (c) includes the spine. (b) Removing pre-segmented bones from the dataset allows to correctly segment arterial branches (d).

These concepts are applied for multi organ segmentation using the pipeline and the software tool described in next section.

3.4 ENDOCAS SEGMENTATION PIPELINE

Starting with the previous ideas and from experiences done in the elaboration of MDCT datasets using general purpose segmentation tools, a functional pipeline, through close collaboration between engineers and radiologists, has been developed.
The experiences done suggest to apply, for each organ, a specific segmentation sequence that consists of the actions reported in Tab 1.

a)	Bounding box selection, to reduce computational time. This action can be manual but however it could be very simple and fast.
b)	Removing pre-extracted structures from the dataset, to take advantages like that shown in figure Fig. 10. Pre-extracted structures can be dilated in order to take advantages in cases of non perfect organs alignment between different phases. This action could be completely automatic.
c)	Anisotropic filtering (optional), to enhance image quality (specially required for obese patients).This action could be completely automatic.
d)	Tuning and running the segmentation algorithm.
e)	Filtering and hole filling, to enhance the quality of the result in terms of mesh smoothing and complexity. This action could be completely automatic.

Tab 1 *General actions of the EndoCAS segmentation pipeline.*

Today there are no fully automatic algorithms usable in action *d* for the segmentation of abdominal structures. In fact full automatic approaches based on anatomical atlas cannot be applied due to natural anatomical variants and to the necessity to work with anatomies modified by interventions and other techniques, based on topology free approaches, work only with few organs [31].
So for action *d* of Tab 1 we can adopt only semi automatic techniques, where the user set input parameters, run the segmentation, watch the result and modify input parameter up to obtain a correct result.
An important aspect during this steering parameters process is the possibility to watch the 3D model obtained with the actual setting. This is more useful than watching only “coloured

slices" (like in almost all general purpose segmentation tools), because the radiologist can validate immediately on the basis of his/her anatomical knowledge the topology of the segmented structure and he/she can correct quickly erroneous results steering the parameters of the segmentation algorithm [32]. For this reason the segmentation pipeline has been implemented as a tool for the open source software ITK-SNAP 1.5.0 by Cognita Corporation *(www.itksnap.org)*, which allows 3D models building, by means of marching cubes meshing, and their visualising easily and rapidly.

3.4.1 Software implementation

The segmentation pipeline has been developed as a tool for the software ITK-SNAP, which is based on ITK and VTK libraries, offers many input and output file formats, standard manual segmentation tools and a semi-automatic snake algorithm. An approach based on region growing has been preferred with respect to the snake and in general to the big family of segmentation techniques based on active contours, because of some practical motivations:

- ✓ application of active contours methods generally requires high mathematical background and the setting of non-intuitive parameters;
- ✓ active contour methods have no objective stop criteria so, the result, also with the same starting parameters, depends on the stop instant given by the user;
- ✓ the total segmentation time would be too long (due to the computational time required by the evolution of the contour), and interactive tuning of parameters could be difficult.

Region growing algorithms, starting from a seed voxel, grow to new voxels (or pixel in 2D case) if they satisfied a defined homogeneity function. Region growing with homogeneity function that take into account the intensity of the neighbourhood voxels (neighborhood connected region growing) [33] offer another advantage. This algorithm includes the current voxel in the object to extract, if all its neighbourhoods, depending on the chose radius, have intensity in the threshold interval (Fig. 11).

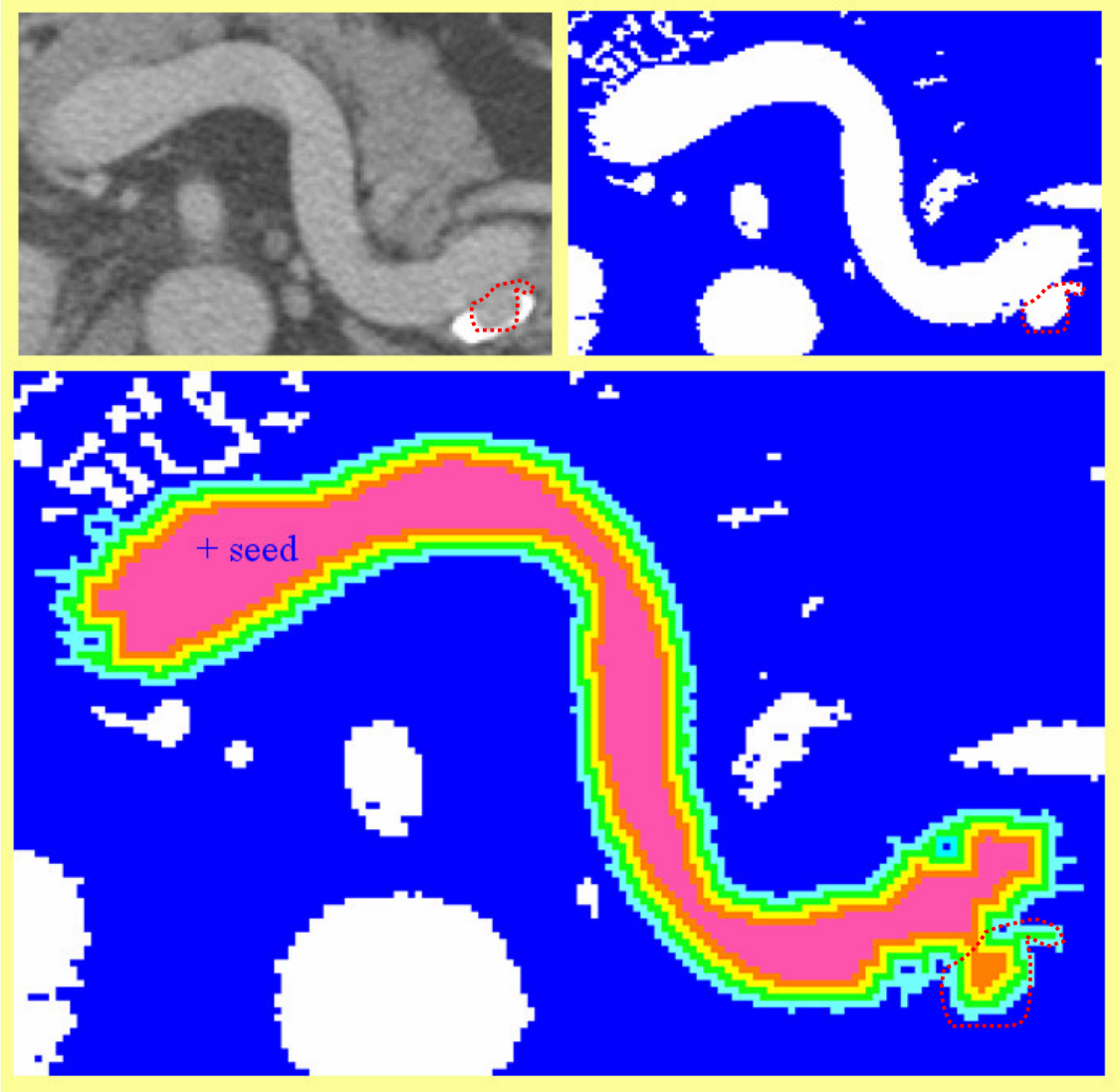

Fig. 11 2D Neighborhood connected region growing example. (upper-left) Original gray-level image. (upper-right) Threshold binarized image. (low) Different result changing the radius. Turquoise region contains pixels four-connected with the seed. In fact for reaching a blue pixel there are others 4 pixels to cross. Reducing the radius other neighbour pixels are included, up to heavenly border, which contain zero-connected pixels. Note that the use of the radius allows to avoid the area inside the red points.

Neighborhood connected region growing allows avoiding erroneous growing trough small vessels or channels created by devices artefacts that join adjacent, but distinct, anatomical structures. For example in Fig. 11 we can observe that using a radius of 4 pixels (turquoise area) we can avoid to include the adjacent splenic artery (inside the red points) and so we can obtain the correct segmentation of the splenic vein.

The tuning of parameters of this algorithm requires just to set a seed point, two thresholds and a radius. The computational time is lower than active contours so the iterative adjustment

of parameters is very fast: if the region grows outside the frontier of the structure to segment, the user can reduce threshold interval or increase the radius and repeat the segmentation.
Fig. 12 shows that using neighborhood connected region growing, passing from a radius of 0 (left) to a radius of 1 (center), a kidney can be discriminated without including renal vein. Using this algorithm the volume of the extracted object will be thinner than the real (depending on the radius), but this error can be compensated by means of an dilatation of the volume (Fig. 12 right).

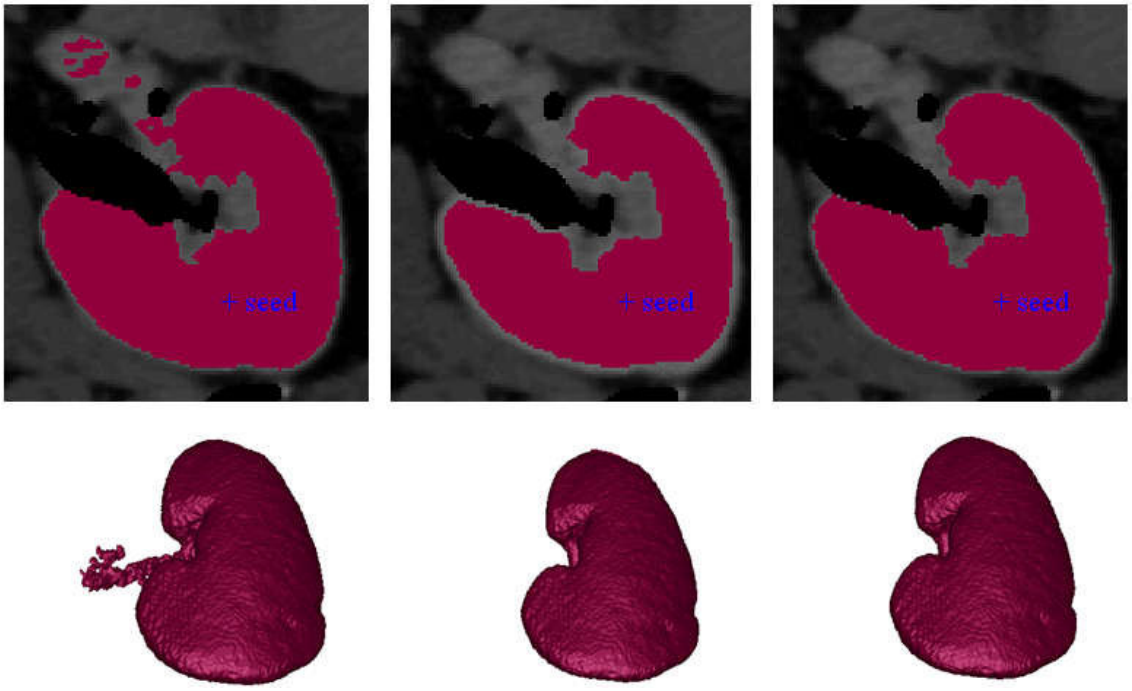

Fig. 12 Result of the segmentation of a kidney from the venous phase, using the same threshold interval, changing the radius and the expansion factor. (Left) radius=0, dilatation=0. (Center) radius=1, dilatation=0. (Right) radius=1, dilatation=1.

Region growing has a stop criteria: "no more voxel to add that satisfy the homogeneity function". It allows to select standard parameters with the certainty to obtain the correct result, without the necessity to stop grow up in a precise instant, at least for some structures easy to segment, like bones (while for others structures the parameters have to be adjusted case by case).
On the previous consideration EndoCAS Segmentation Pipeline has been developed as a tool integrated in ITK-SNAP that it is composed of some steps. An example of the pipeline initialization, that enclose *a*, *b* and *c* actions of Tab 1, is shown in Fig. 13 :

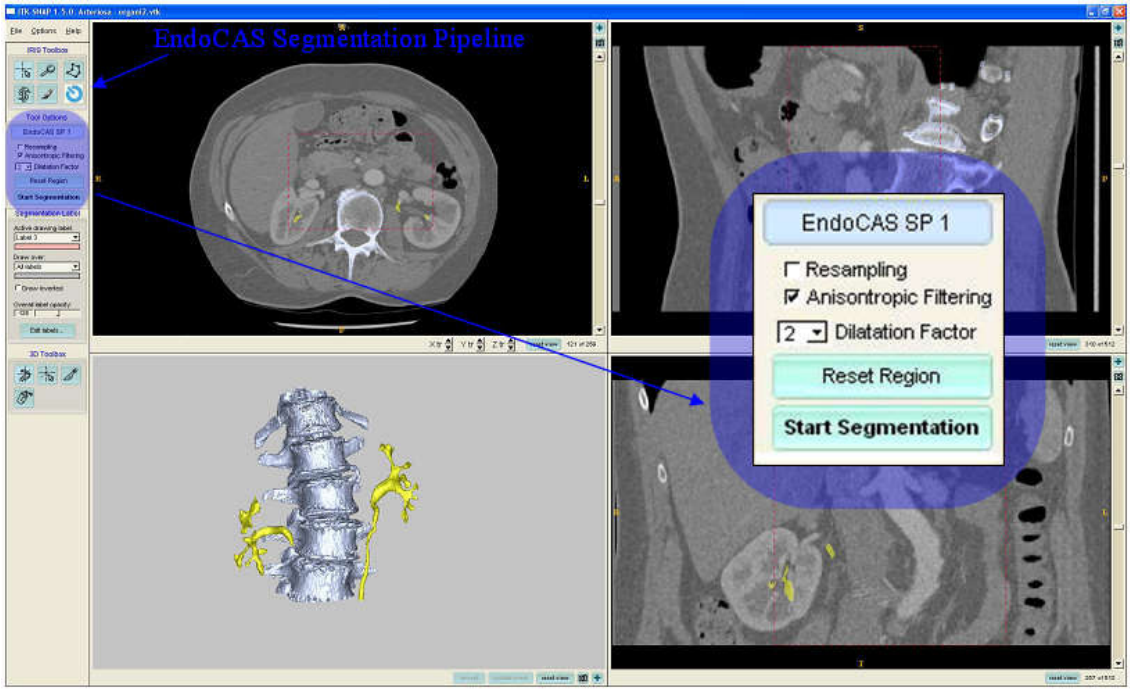

Fig. 13 Initialization of EndoCAS segmentation pipeline in ITK-SNAP. The user selects the working volume, choose the dilatation factor for removing (lumbar spine and ureteres in this case) and select or not image anisotropic filtering. The user can also resample the image, which can be useful to reduce the computational time.

The user can choose to use anisotropic filtering [34] [33] on the selected region, in order to increase the image quality (especially required for fat patients).

The pipeline, after the initialization, proceeds with 3 steps that allow to segment the selected region (*d* action of Tab 1). In Fig. 14 it is shown the application of these 3 steps for the segmentation of the abdominal aorta in the region selected in the previous figure. After the setting of the threshold interval (Fig. 14-a) and the placement of a seed voxel (Fig. 14-b), it is shown the result of the segmentation with 2 different radiuses and expansion factors (Fig. 14-c and Fig. 14-d). Note that the use of the radius, like previously described, allow to avoid erroneous growing in the left renal vein and in the kidneys (compare Fig. 14-c with Fig. 14-d).

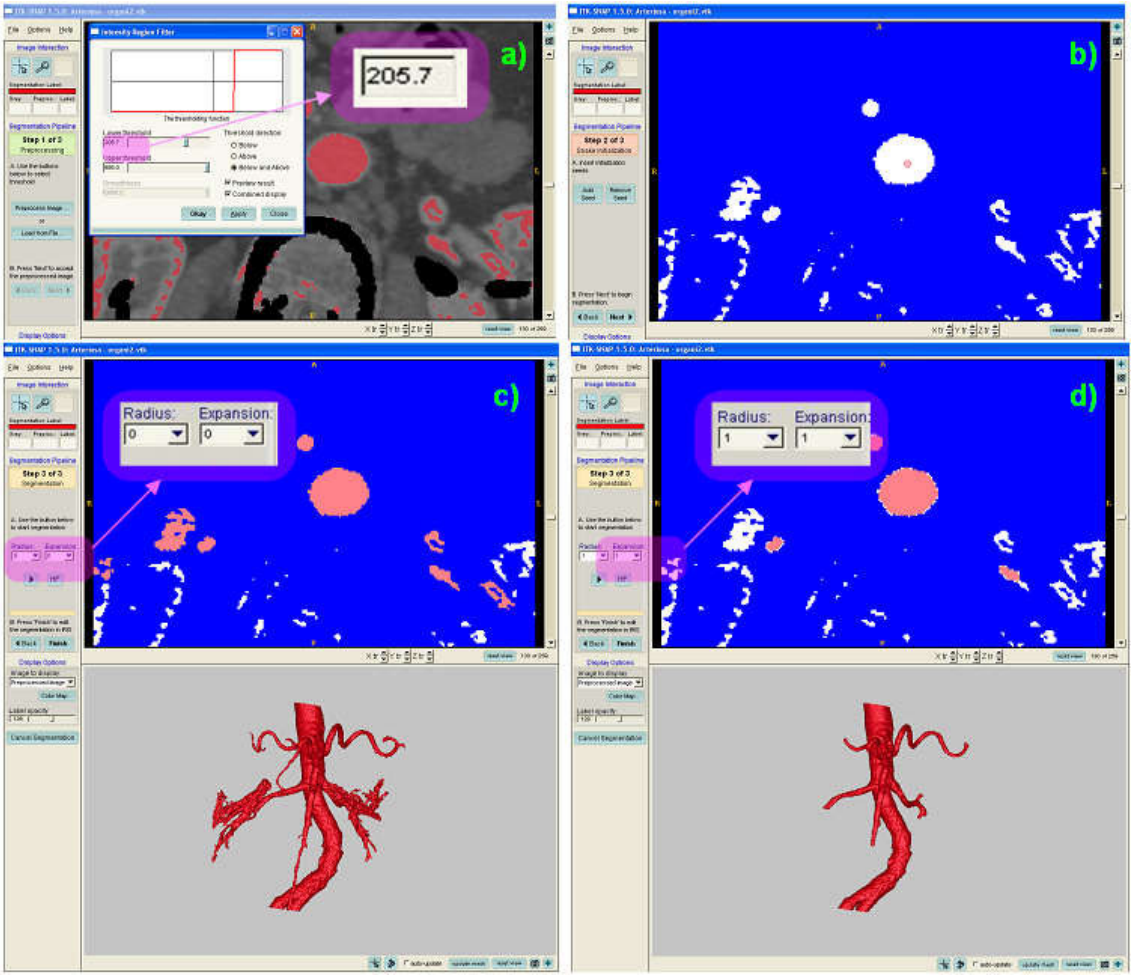

Fig. 14 Working with EndoCAS segmentation pipeline. (a) Step 1: selection of threshold interval (with sliders) seeing in real time the obtained binarized image in the 3 cross section windows (it is shown only the axial section). (b) Step 2: insertion of one or more seeds in the binarized image (white for pixels inside the threshold and blue for pixels outside). (c) Step 3: running of the segmentation with 2D and 3D result. (d) Step 4: result using different radius and expansion factor.

For this abdominal aorta segmentation, like shown in Fig. 15, it is better reducing the threshold interval that increasing the radius. In fact increasing the lower threshold we obtain, with radius=0, much parts of the aorta sub-vessels (compare Fig. 15-b with Fig. 14-d). In general this approach (reducing the threshold interval before to try to use the radius) offers best results, because the use of the radius does not allow to include small vessels or small pieces of organs.

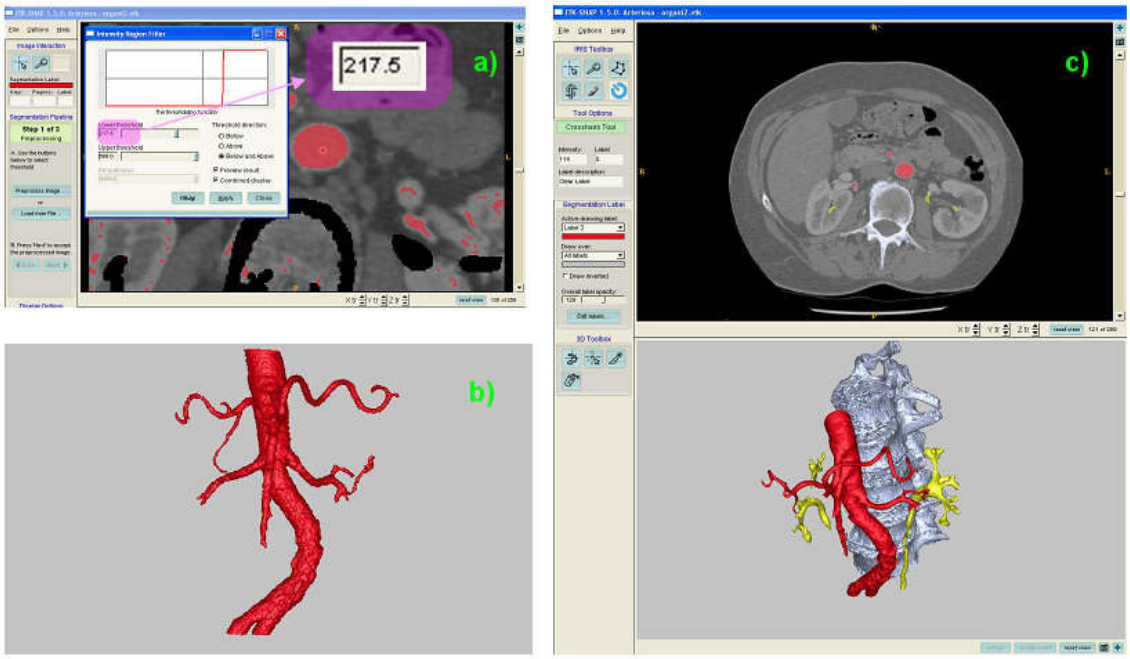

Fig. 15 Change of the threshold interval in EndoCAS segmentation pipeline. (a) Increasing of the lower threshold respect to the previous figure. (b) 3D model obtained with the new threshold interval and radius = 0 (c) Principal ITK-SNAP window with the just extracted abdominal aorta joined to ureters and the spine (previously segmented).

Filtering (first part of the action *e* in Tab 1) is offered directly by ITK-SNAP during the generation of 3D models. The user can set parameters of a gaussian smoothing that is applied on the segmented dataset before the generation of the surface. ITK-SNAP has also decimation and smoothing filtering to apply to the surface generated.

Hole filling (second part of the action *e* in Tab 1) is offered to the user in step 3 of EndoCAS segmentation pipeline, where the user, after the segmentation, can apply the filling algorithm pressing a button. This algorithm applies a region growing on the segmented binary dataset, with a seed positioned outside of the structure of interest, and after it takes the complementary of the result. Note that this algorithm removes all holes, so it has not to apply if the anatomy to extract has real internal holes.

In order to compensate variations of the patient setup, during the acquisition of the multi-phase datasets, 3 buttons have been introduced in ITK-SNAP for manual rigid registration (only translations) under visual control, that allow to align each phase to the current segmented image.

Summary of all steps of the pipeline are reported in Fig. 16.

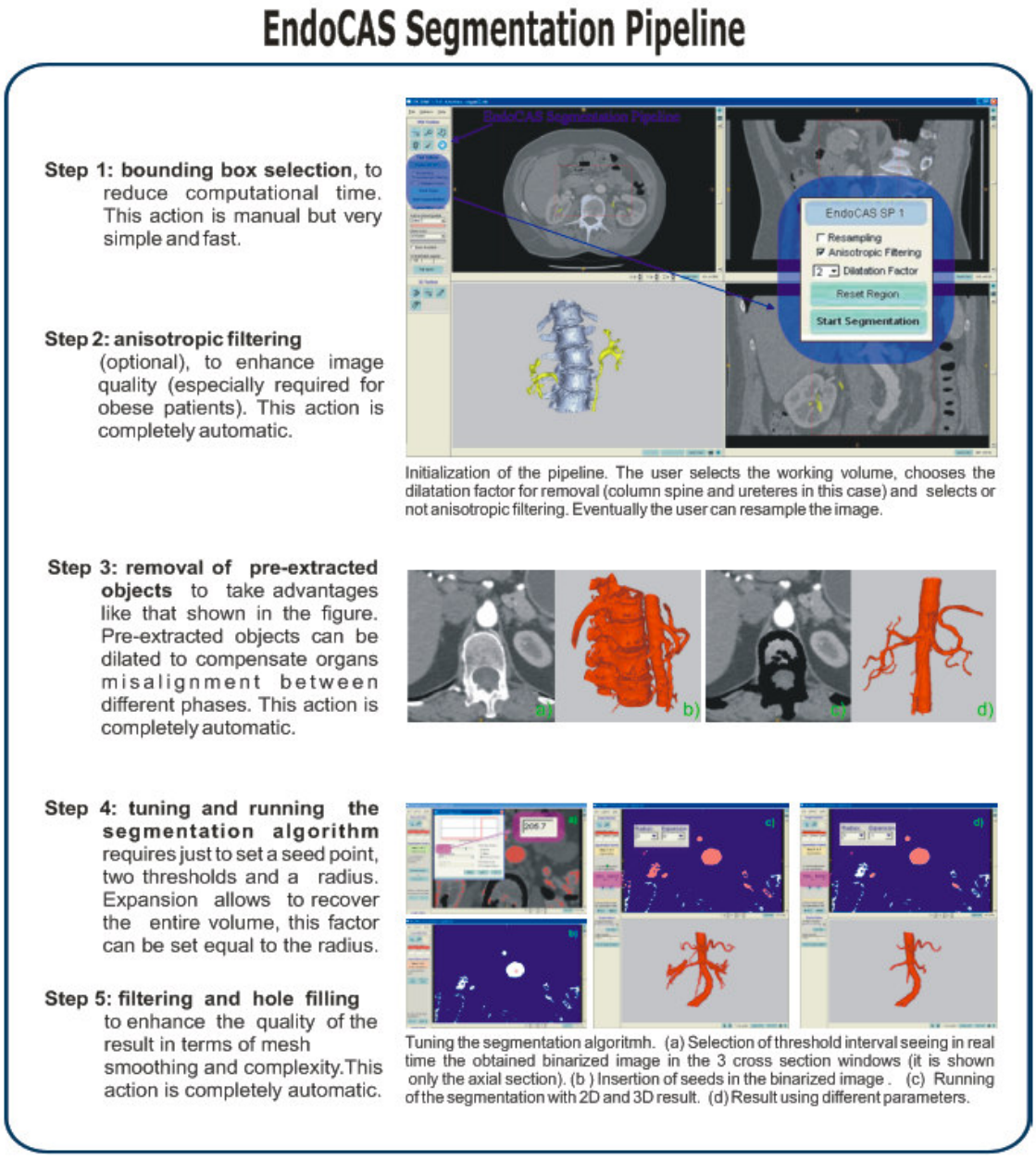

Fig. 16 Schematic representation of the steps of EndoCAS segmentation tool integrated in ITK-SNAP.

3.5 OPTIMAL SEGMENTATION SEQUENCE

As previously introduced, the difficult for the segmentation of an object is reduced removing pre-segmented objects, and the sequence of objects extraction play a fundamental role.

An optimal segmentation sequence has been defined on the basis of anatomical, empirical and functional consideration concerning the timing of contrast distribution. This sequence

can be used with MDCT datasets commonly used for diagnostic purposes. This segmentation sequence requires an urographic phase for the segmentation of ureters, while, for others structures, arterial, venous and basal phases are enough. Obviously the availability of more phases can simplify and speed up the segmentation procedure. The table in Fig. 17 shows the optimal segmentation sequence that is motivated below.

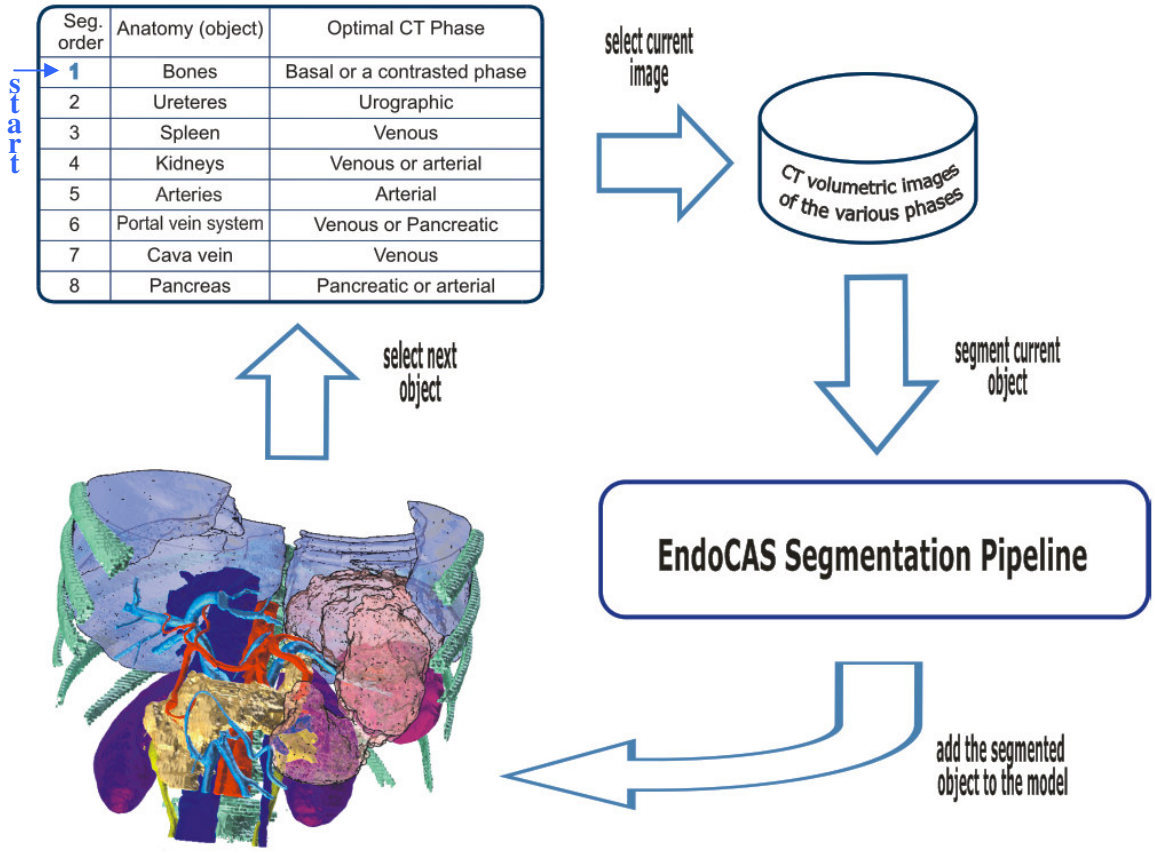

Fig. 17 Optimal segmentation sequence. At first, it is selected the optimal phase for the segmentation of the current anatomy and the corresponding volumetric image is loaded for the segmentation. After the segmentation the just extracted object is added automatically to the others.

Bones

The segmentation of spinal column and ribs is useful, further than for their visualization (or other scopes), also for simplifying the extraction of the other structures. For example (as previously described) abdominal aorta is connected to spinal column by lumbar arteries and in some cases uterers are very close to the ribs, so the segmentation of bones and their removing simplifies the extraction of these structures (Fig. 18 right).

Bones are the structures with the biggest intensity in a CT exam, so they can be segmented at the beginning. If a basal phase is present, spinal column and ribs can be segmented from it, using just connected region growing (radius=0) without anisotropic filtering. For spine segmentation is needed in general only one seed on its centre. In order to obtain a surface with better quality, hole filling has to be applied. Often the basal phase have the biggest thickness of the exam, so, if it is required a detailed model of bones, they can be segmented from a contrasted phase but taking into account that often it is necessary the use of the radius (of the neighbourhood connected region growing).

Obviously if in the exam some ribs are not connected with the spine, due to the volume selected from the radiological technician, like in Fig. 18 left, a seed for each non connected rib is necessary.

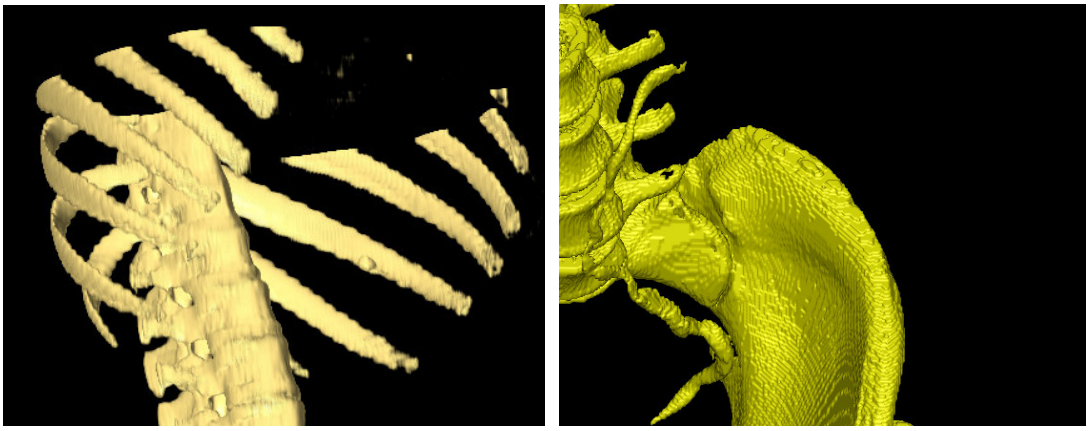

Fig. 18 – (left) Spine segmented from the basal phase: is necessary to insert a seed for each isolated parts. (Right) Bad ureter segmentation from urographic phase: sometimes the basin is very close to urinary system, so, without its removing, a seed positioned in the uretere can grow up to the spine.

Ureteres

Ureteres can be extracted from the urographic phase, using just connected region growing (radius = 0) without anisotropic filtering. If an ureter is not full of urine a seed for each visible segment is needed. Furthermore in few cases the ureteres can touch the basin (see Fig. 18 right); in that case removing of bones is required.

Spleen

The segmentation of the spleen is generally easy from venous phase (Fig. 19 left). Here it's less contrasted than the vessels and more contrasted than the structures very close to it. It is preferable to extract the spleen before the kidneys because it simplify the discrimination of

them. The segmentation require smoothing with anisotropic filtering, neighbourhood connected region growing (in general radius=1 voxel is sufficient) and holes filling. Note that the lower and the upper threshold has to be very close.

Kidneys

The segmentation of kidneys is very easy from the venous phase if it is fine contrasted. In this case is enough anisotropic filtering, just connected region growing and eventually holes filling (Fig. 19 right).

If the acquisition of the venous phase is too premature, the interface between the kidney parenchyma with its vein is difficult. In this case the user can try to use the arterial phase. In fact the contrast flows at first in the artery, after that in the kidneys and then in the veins. So the extraction of kidneys is possible also from the arterial phase, but the discrimination of the interface, between kidney and artery, often requires the use of the radius.

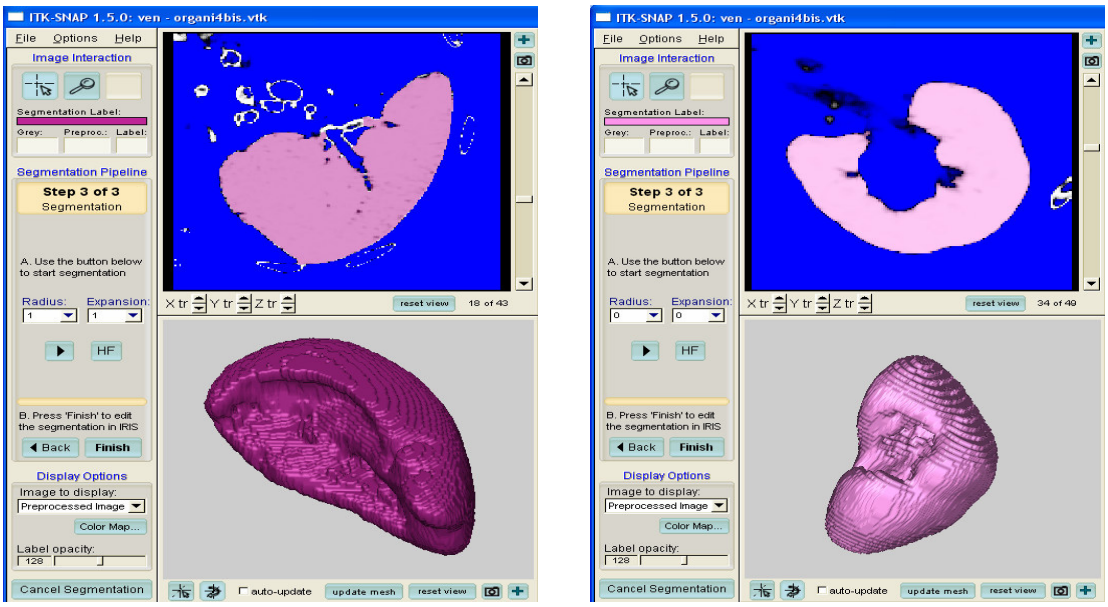

Fig. 19 - (Left) Segmentation of the spleen from venous phase using radius = 1. (Right) Segmentation of a kidney from venous phase without using of the radius.

Arterial vessels

The segmentation of the aorta with all sub-vessels could be difficult, but thanks to the removing of the other structures can be performed using connected region growing and holes filling. The anisotropic filtering is required if the image is noisy, due to an erroneous contrast

timing or in presence of obese patients. Some times, especially if the arterial phase is delayed, the discrimination of the interface artery/organs requires the use of the radius for their discrimination.

Splenic, mesenteric and portal vein

In this case the segmentation can be done using anisotropic filtering and neighbourhood connected region growing from venous phase (in general radius=1 is enough). Some times it is necessary more than one seed in order to extract parts of the vessels, which can be isolated due to the using the radius.

If it is present a pancreatic phase (or a portal phase) the work can be simplified. The contrast medium flows before in the portal vein system and after in the cava vein. Therefore, splenic, mesenteric and portal vein segmentation can be simplified using pancreatic phase, because their discrimination whit the cava vein (difficult in some cases using the venous phase) is more simple.

Cava vein

Cava vein can be extracted from venous phase using anisotropic filtering and using the radius. Unfortunately cava vein is not recognizable in the track inside the liver, so box selection has to be used in order to avoid liver parenchyma and vessels. The inclusion of liver segmentation in the sequence would simplify cava vein segmentation. It has been tested segmenting the liver like shown in Fig. 22.

Pancreas

The segmentation of the pancreas is not easy in none phase. Its segmentation requires great attention and experience by the user. However, if the user is trained, anisotropic filtering and neighbourhood connected threshold with radius = 2 allow, in general, to extract the pancreas.

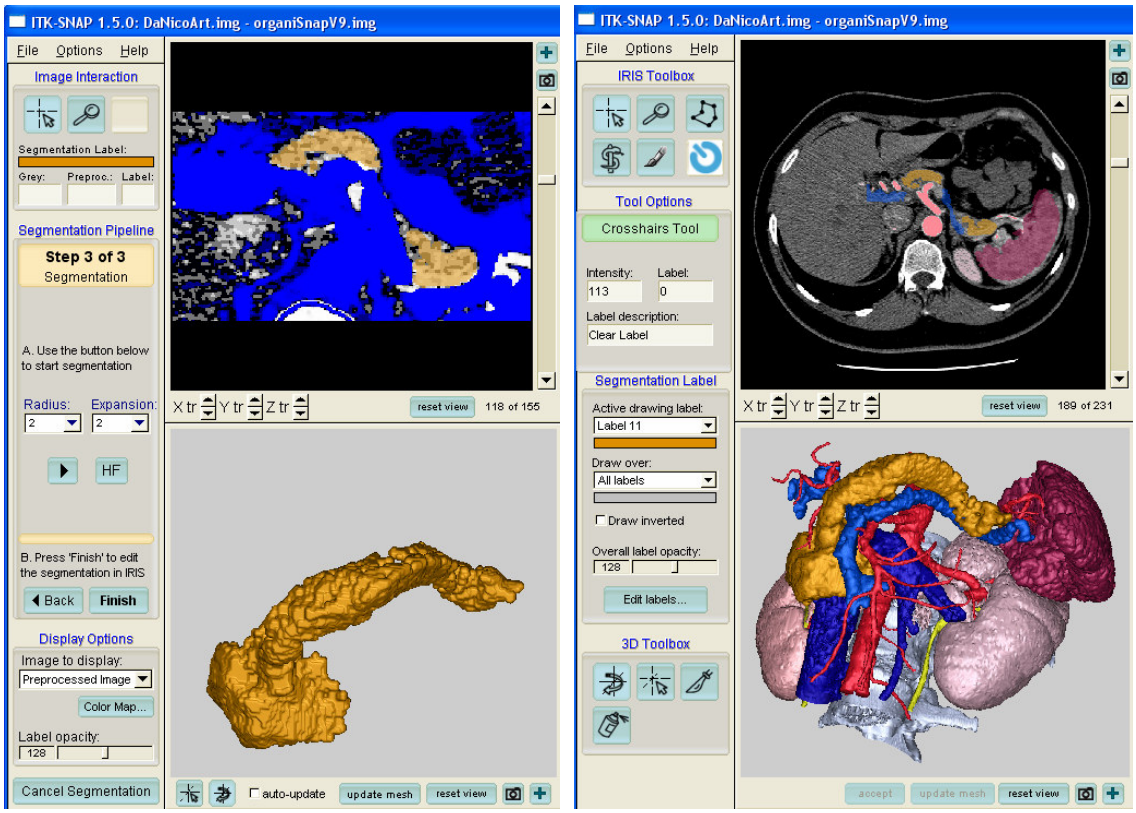

Fig. 20 - (left) Pancreas segmentation from arterial phase using the radius. (right) Complete model of the same patient.

The last 3 anatomies of the segmentation sequence are the most difficult, however, if the user is trained, and the images have enough quality, the extraction can be done repeating 3/4 times the segmentation whit different parameters (for each object).

Finally a note on the dilatation factor for the pre-segmented structures (automatically removed from the current image), which depends on the misalignment between the phases of the dataset (low with aligned phases, high with not aligned phases).

3.6 IMAGES ACQUISITION PROTOCOL OPTIMIZATIONS

The segmentation procedure requires that arterial, pancreatic and venous phases are acquired with slice thickness at least of 2.50 mm (1.25 mm recommended). Otherwise all small vessels will be not reconstructed ("followed") by the region growing algorithm. Regarding basal phase, used only for bones extraction, can be with slice thickness up to 5.00-10.00 mm (if the quality of the mesh for bones is not important).

Even if the segmentation procedure proposed can work with all images used for traditionally diagnostic purposes, with enough thickness, some optimizations on the acquisition protocol can simplify the segmentation work.

First of all, the CT technician, in order to optimal empathize the structures involved by the specific phase, have to take a great deal of attention on the acquisition instants and on the contrast injection. For example, it is better to do not enhance organs in the arterial phase, because it extremely simplifies the automatic discrimination of arteries near parenchyma of organs. I noted that, in traditionally protocols, this aspect some times is not empathized, because probably, the radiologist, basing on his/her experience and intelligence, can not exchange vessels for organs, like automatic algorithms sometimes do.

Further voxel positions have to be stable in the various phases as more as possible. It allows guaranteeing spatial relation by anatomical structures and amplifies the effect of the subtraction mechanism (that simplifies the segmentation). In case of misalignment, the images of the various phases can be pre-registered, but it requires additional work. So it is better to ensure a comfortable position for the patient (compatible with the exam to perform) using cushions and immobilization devices, in order to avoid movements, and it is important to inform him/her to stay more stable, as more as possible, not only during the acquisitions, but during all the examination.

A supplementary source of movements for organs is the breathing. For this reason abdominal images are in general acquired with the patient in apnea. In order to guarantee the same position of the organs in each phase, the diaphragm position (that moves organs) can be controlled using breathing monitors or triggers. It was proved that using a simple monitor device, like the one shown in Fig. 21 (developed by the author), we can guarantee the stability of the position of organs at fixed points of the breathing cycle.

Fig. 21 Simple two steps breathing monitor mounted on a phantom. Two magnetic switches (in the transparent box) are mounted at different distances from a calamite (white), on an elastic adaptable belt to fix around the thorax, which turn on/off the red and the blue high emission leds in function of the thorax expansion.

The sensor is based on two magnetic switches, mounted at different distance from a calamite, on an elastic adaptable belt to fix around the thorax. Two coloured high emission leds turn on/off, in function of the thorax expansion. It has been proved, using an US, that training a patient to inhalate up to turn on only one led (the blue one in the specific case), and stay in apnea, we can control the cranio-caudal position of kidneys with an accuracy of 6-7 mm. The useful of this sensor has also been verified in the acquisition of some MDCT contrasted images. The acquisitions have been started when the (trained) patients turn on the first led (the second one designate a too big inhalation). It has been obtained good organs alignment, so, as consequence, the segmentation work was been simple, thanks to an useful subtraction mechanism on the various phases. These results demonstrate the useful of breathing monitors for segmentation using the proposed procedure, so, in the future, it would be very functional to use monitors appositely designed for MDCT scanners (metal free and with a continuous thorax expansion measurement). Regarding breathing triggers, they could command automatically the acquisition, like in the work [35], allowing a perfect alignment in free breathing, so obtaining stable images with not cooperative patients too. In this case a complete scansion of a patient requires the acquisition of some consequential portions of

volume at the same instant of the breathing cycle. It is preferable to start the acquisition at the end of the exhalation phase, because the diaphragm remains stable for a little of time (see Fig. 29 on the next chapter). The number of volume portions depends on the scanner velocity (direct correlate to the number of detectors). Unfortunately there are no scanners today so fast to acquire an abdomen using only one volume during free breathing. For this reason the use of breathing triggers, which would allow free breathing, is difficult in case of contrast agent injection because, in this case, a rapid and "on time" acquisition is required, so no trigger commands can be waited.

Finally the segmentation sequence has to be adjusted in function of the acquisition procedure used in the radiological department, and on the abdominal structures targets of the segmentation work. For example pancreas segmentation can be extremely simplified if the patient drinks water and receives buscopan before the exam. It allows to guarantee the discrimination (in general difficult to obtain) between pancreas and colon. The colon will be full of water and stable (no spasms), so with an intensity very different respect to the pancreas.

3.7 THE SEGMENTATION SOLUTION AT WORK

The whole segmentation procedure is based on the neighbourhood connected region growing algorithm that, appropriately parameterized for the specific anatomy and combined with the optimal segmentation sequence, allows optimal segmentation results in spite of simplicity and usability. This basic algorithm, integrated in the EndoCAS segmentation pipeline, is efficient for clinical purposes, in terms of usability, computational time and quality of the result. Instead to follow the trend to automate completely the segmentation process (a very important challenge for future CAD/CAS systems, encouraged by international competitions), where the physician watches the result only when the segmentation is terminated, this segmentation solution offers an easy and simple method, where the radiologist guides the segmentation process using his/her experience and anatomical knowledge. Interactive visualization of the 3D anatomical model, obtained with current settings, can be used by radiologists for a fine tuning of segmentation parameters, allowing to build 3D models, usable in the clinical practice, today.

In addiction all 3D anatomical models have to be validated before their clinical using. In the EndoCAS segmentation pipeline the validation is made by the radiologist directly at the workstation, during the segmentation, steering the procedure up to obtain the desiderated result. Regarding validation is also evident that a radiologist that see the growing of each virtual anatomies, one by one, take implicitly more attention to the global result, respect to a radiologist that has to validate a segmentation done by an automatic algorithm, sliding entirely a "coloured" volume superimposed to gray scale images.

Obviously the proposed segmentation approach can not be used with datasets acquired without finalized protocols, for example with elevated slice thickness, too poor quality or non contrasted images, but the intention is to work with images that came from normal radiological practice, where it obtain good results.

The segmentation solution proposed allows a radiologist to generate and to validate 3D models of bones, ureters, kidneys, pancreas, spleen, arterial, portal-splenic-mesenteric and cava veins in about 30 minutes of work. It has been tested that a new user, after some hours of training, can use the software proficiently. 3D models have been generated for the planning of some interventions and in the EndoCAS Navigator for a laparoscopic distal pancreatectomy performed in the Cisanello Hospital. In the future, the automation of some manual activities could reduce the total segmentation time up to 15 minutes, so a radiologist could build 3D models of each patient, during the diagnosis of each exam.

3.8 ADDITIONAL FUNCTIONALLITIES DEVELOPED FOR THE SEGMENTATION SOFTWARE

ITK-SNAP software has been integrated, over than with the EndoCAS segmentation pipeline, with additional functionalities. First, in order to work with multi-phase datasets, containing images with different dimensions, in terms of volume acquired and in terms of voxels resolution (a very frequent situation in the clinical scenario), the image loader has been modified. ITK-SNAP stores segmentation data in volumetric images with a format equivalent to the format of gray scale radiological images (Dicom, VTK, etc.) containing voxel and volume information. If the dimension of the image that contain the segmentation is not equal to the dimension of the gray scale image, the modified image loader performs an alignment and an adaptation in order to obtain a bi-univocal correspondence between vovels

of the segmentation image and voxels of the gray scale image. It allows to segment some objects from a phase and after that to change the phase using the previous segmentation image, without manual work for image adaptation.
The others functionalities are described in the next two paragraphs.

3.8.1 Object volume calculator

This tool has been developed in answer to a surgeon's requirement: volume evaluation for anatomical regions. These measurements are important for the estimation of the extension of diseases or for the evaluation of remaining functionalities after tissue removal (see Fig. 22).

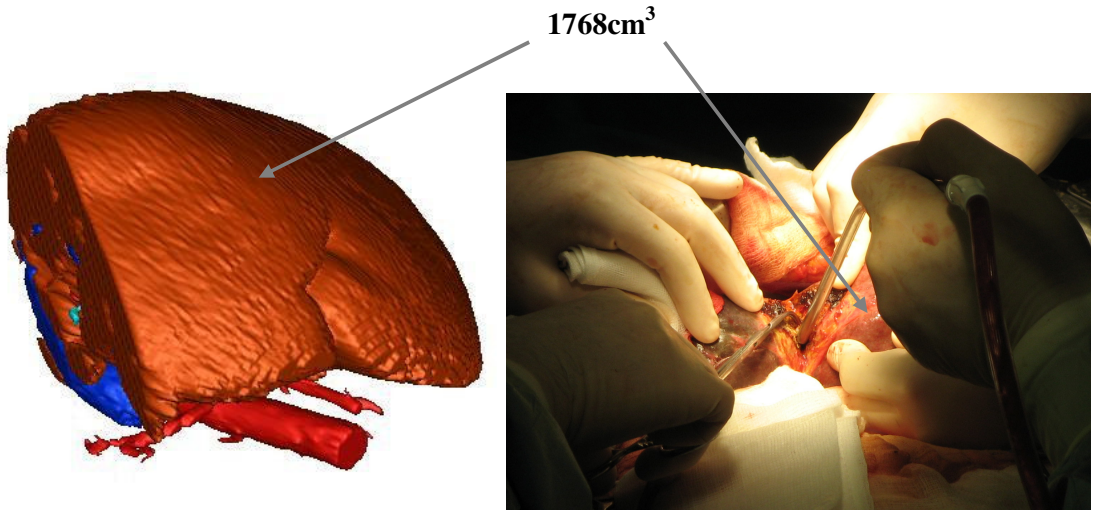

Fig. 22 (Left) The volume calculator integrated in the ITK-SNAP allows to evaluate the volume of each segmented region. The image is given from a real case, in the planning phase before a right epatectomy, where the surgeon cut the liver (using the virtual cutter tool) and evaluated the volume of the remaining liver. (Right) Surgeons at work during the epatectomy.

Having voxel dimensions, the implementation has been simple. The number of voxels contained in a region are multiplied for the volume of a voxel. Obviously the real volume and the virtual volume are identical only if the segmentation is perfect, otherwise their difference follows the percentage of the wrong segmented areas.

3.8.2 Drilling guide for dental implantology

As reported in the thesis introduction, drilling guides are useful in dental implantology for surgeon assistance in drill positioning. Watching the computer assisted implantology workflow, the author seen that ITK-SNAP software, added with few modifications, could be used for this purpose.

In general, drilling guides are composed of a plastic model that fits on the mouth with encapsulated cylindrical metal bushes for drilling guidance. A first mask model has been realized segmenting the volume around the teethes and the mouth mucosa of a CT real image. The model (exported in STL format) has been realized using a stereolitographic printer and tried in the patient mouth. Subsequently to fitting results, validated positively by a surgeon, a tool for cylinder positioning has been integrated in ITK-SNAP. Like shown in Fig. 23 the surgeon can chose the plant axes, setting two 3D points in the volume, after that he/her can choose internal radius (in function of the bush to fix inside), external radius and the length.

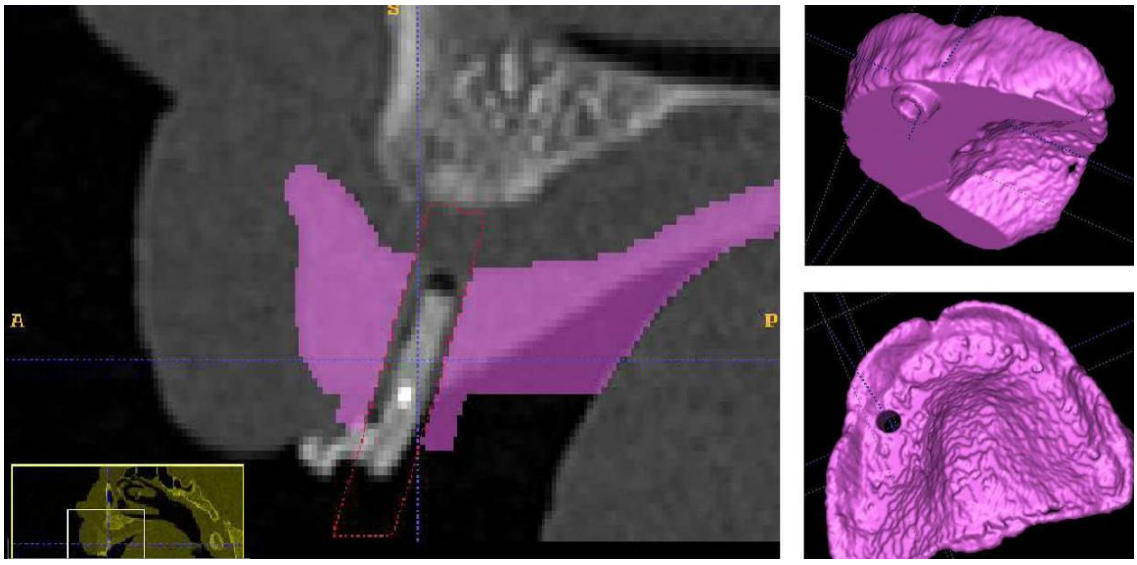

Fig. 23 Dental implantology tool. (left) Dental plant axes positioning on the CT images and cylinder drilling guide dimensions setting. (Right) Rendering of the surgical guide with the hole for the metal bush.

4.0 ABDOMINAL CAVITY REGISTRATION

The purpose of this chapter is the evaluation of error sources that influence the registration of pre-operative information to the real patient, and the discussion of a methodology to employ, in order to reduce inaccuracies and allow the introduction of computer guided surgery for the abdominal district. It is reported some suggestions for the preliminary rigid body registration to use directly in the surgical room or to use as starting point for more accurate deformable approaches.

4.1 CONCISE STATE OF THE ART

One of the not completely resolved key functions, for building CAS systems [14], is the robust registration of the virtual pre-operative information (acquired by means of radiological devices) with real intra-operative information (consisting of the patient on the surgical table). This task requires to determine the geometrical transformation of correspondent points taken in two different reference frames and in two different time instants. In fact, pre-operative information is given in the reference frame of the radiological device and it is acquired some days before the intervention, while the intra-operative information is referenced to a reference frame of the surgical room (in general defined by means of a tracking system) during the intervention.

Many approach deriving from general purpose image rigid and non-rigid registration methods [36] are adopted for medical images [20]. In addiction to these analytic works, many computational intelligence methods were proposed [37]. All these works solve a general problem of registration between two dataset compensating different positioning of the patient, device aberration and in some cases anatomy shifts [38]. Rigid registration is used for many CAS applications, and various analytic and experimental accuracy evaluation works have been developed, but for the complexity of the dynamics of abdominal and torso structures, these applications are limited to bones and head. Non rigid registration algorithms, like for example [39], can offer good results, but their using for CAS applications would require dense and volumetric intra-operative information (like open CT or MRI).

For the abdomen, like for all deformable structures, registration could be eventually skipped using intra-operative imaging [40]. A list of intra-operative registration free devices/applications can be found in [41]. These approaches are intrinsically very precise concerning voxel localization, but, unfortunately, their visual quality is limited by the poor definition of intra-operative radiological scanners compared with the quality of pre-operative scanners and also this methods intrinsically require a volumetric imaging device in the surgical room, that in general is not present.

As shown in [42] trajectory of target points could be predicted, during breathing motion, at given respiratory states. Also recent works [43] [44] propose dynamic motion models that reproduce the entire trajectory during the breathing cycle. In these works are compensated physiological movements due to breathing, but they can still fail for movements generated by other physiological effects or different decubit of the patient on the surgical bed.

For a complete dynamic registration that compensate all physiological parameters and different decubit of the patient in the surgical room relative to that one assumed by the patient in the radiological department, we would have an exhaustive biomechanical model of the anatomy like those, foretold in [45], tried in [46] and encouraged by the EU [47] .

We can note that all non rigid techniques, static or dynamic, require a preliminary coarse (less possible) registration to use as starting point for more accurate deformable approaches [20] [48]. Today, rigid-body registration is the unique approach used for this initial alignment owing to the amount of works published that evaluated its mathematical properties and for the existence of optimal solutions for the problem. Even if there are some studies about the accuracy of rigid-body registration for cranial and extra cranial targets [49], there are no apposite works for the estimation of errors with this type of registration on specific abdominal structures.

4.2 THE ABDOME REGISTRATION PROBLEM

There are many parameters that influence the localization of targets in the surgical room using their position acquired with radiological devices before the intervention [50]. Like previously described, for some structures, like bones, their position depends only by the steady pose initially assumed by the patient on the surgical bed, but for other non rigid objects, like the heart, the lungs and the abdomen, their position change dynamically.

In this chapter Vrd_i identifies the position of the centre of mass of a generic voxel *i* acquired in the radiological department by means a scanner with its reference frame. V$sr_i(t)$ is the trajectory of the point corresponding to the center of mass of the same voxel in the surgical room, during the intervention, relative to a reference frame (generally a tracking system) mounted in the surgical room (see Fig. 24). The optimum for surgery would be to know exactly where is each $Vsr_i(t)$ in each instant time, with extreme precision. In the rest of the chapter $\hat{V}sr_i(t)$ identifies the estimation of V$sr_i(t)$, in the reference frame of the surgical room.

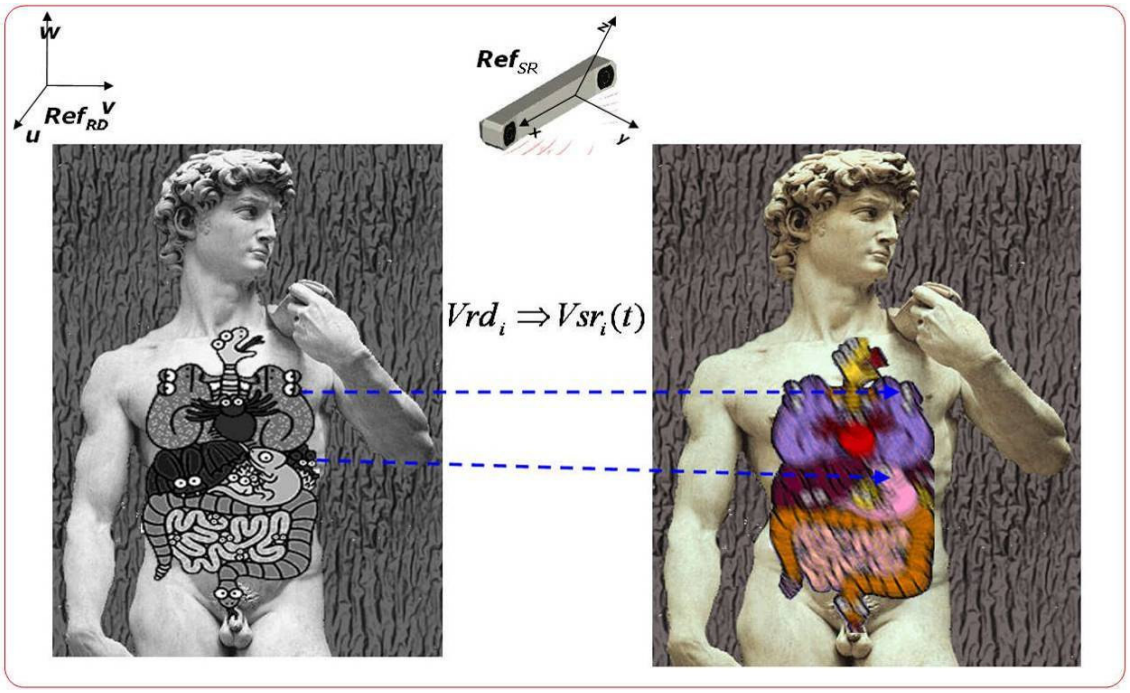

Fig. 24 Voxel baricenter Vrd_i statically acquired in the radiological department (left), has to be remap on the moving corresponding V$sr_i(t)$*(right)*[1].

A similar problem is solved every day by radiotherapists. The difference is that they have to localize only a restricted target area and not an entire anatomical district, like in general for CAS, and also they have all reference frames in the planning room (normally a CT with a positioning system based on laser crosses) mechanically calibrated with all reference frames in the bunker room (normally a proton accelerator, a positioning system based on laser

[1] I would like to thank the unknown author of the funny anatomy depicted in the figure.

crosses, and an eventually a 2D imaging device). In that case they do not have to register the patient, but to reposition him/her in order to coupled laser crosses with reference markers on the skin of patient skin.

4.2.1 Factors that influence the registration

At first a change in bed position, in terms of angles relative to the floor, determines a movement (fall down) of internal organs. Also the shape of the bed has to be taken into account because it introduces a deformation: further the (obvious) case of bendable bed, also concave beds (like some used with CT and MRI devices) introduce a deformation on the abdominal cavity. Furthermore the abdominal cavity is delimited by the basin (a rigid object) and by the thoracic case (a semi rigid object), so, the relative positioning of these anatomical structures, depending by the patient decubit, influences the location of the contained anatomies. The configuration assumed by the limbs can change the position of abdominal structures too.

After that we have to take into account errors introduced by the radiological device and in all post processing phases (like segmentation, for example). Also the use of particular radiological devices or particular surgical techniques imposes some requirements on the patient that, as reflect, introduce errors in the localization of pre-operative targets. For example, during some abdominal interventions, an arm of the patient lies along his/her torso, so it would be advantageous to acquire the patient in this position, but CT scansion of hands, close to the abdomen, introduces artifacts in the image, so hands are (generally) positioned in the rear of the head, so introducing a variation in the patient setup. Furthermore, generally, the scansion with CT devices requires that the patient inhales a quantity of oxygen in order to stay in apnea during the acquisition, with the purpose of acquiring as more as possible a steady volume. The effect of a great inhalation introduces a fall down in the diaphragm and also in the organs in contact with the diaphragm itself. This fact is used every day by radiologists in US acquisition, to better expose organs during examination, which know that organs can be moved up to 6-8 cm (tried in the lab), using breathing. This effect has been also evaluated on a 65 years old woman, aligning her spine on a CT dataset acquired during normal breathing to a CT dataset acquired during inhalation apnea. The result in Fig. 25 shows that kidneys have been moved up to 27 mm, while the displacement of the spleen has been evaluated up to 40 mm.

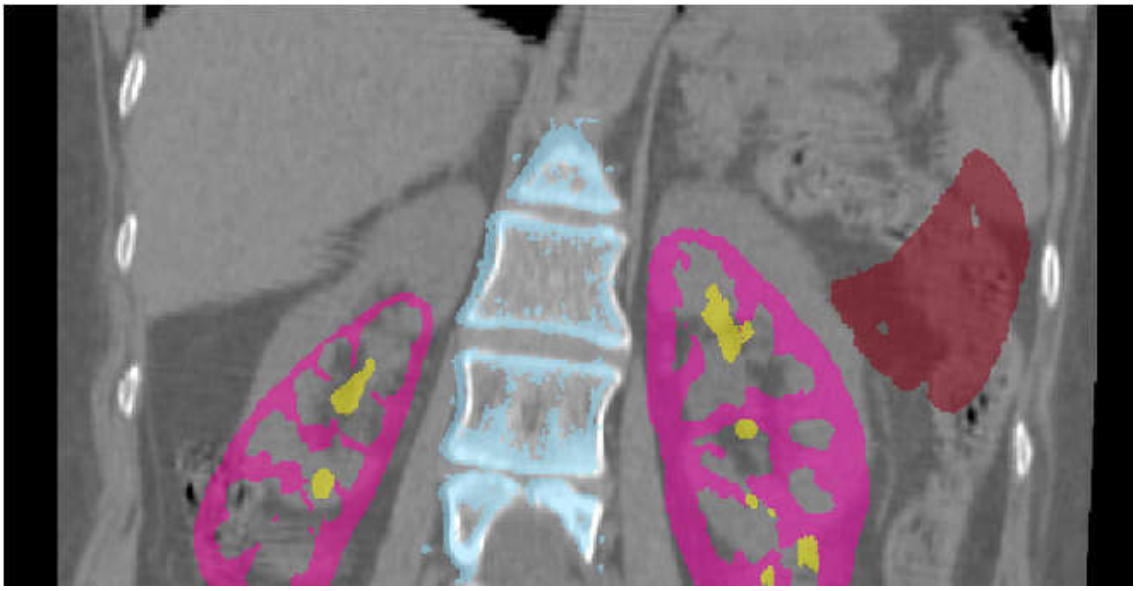

Fig. 25 The spine of a patient, acquired during normal breathing, is aligned to a dataset (segmented and coloured) acquired during inhalation apnea. It can be evaluated the displacement of spleen and kidneys.

Also normal breathing introduces a displacement and a deformation (especially) on upper abdominal organs. The diaphragm has a maximum excursion during normal breathing of about ±7 mm [51] around its median position, so, as consequence, the structures in direct contact with it would have a similar excursion and deformation. The experiment described in the next paragraph shown that the abdominal surface has a maximum excursion of ±7.5 mm during normal breathing. The other internal organs, passive elements enclosed between the diaphragm and the abdominal wall, would not move (probably) more than them.

Another important class of movements is caused by the gastroenteric system. Filling level and spasmodic movements of the stomach and the intestine determine a displacement for structures in the neighbour of them. These errors are particularly relevant on intra peritoneal structures. See Fig 26 for an example.

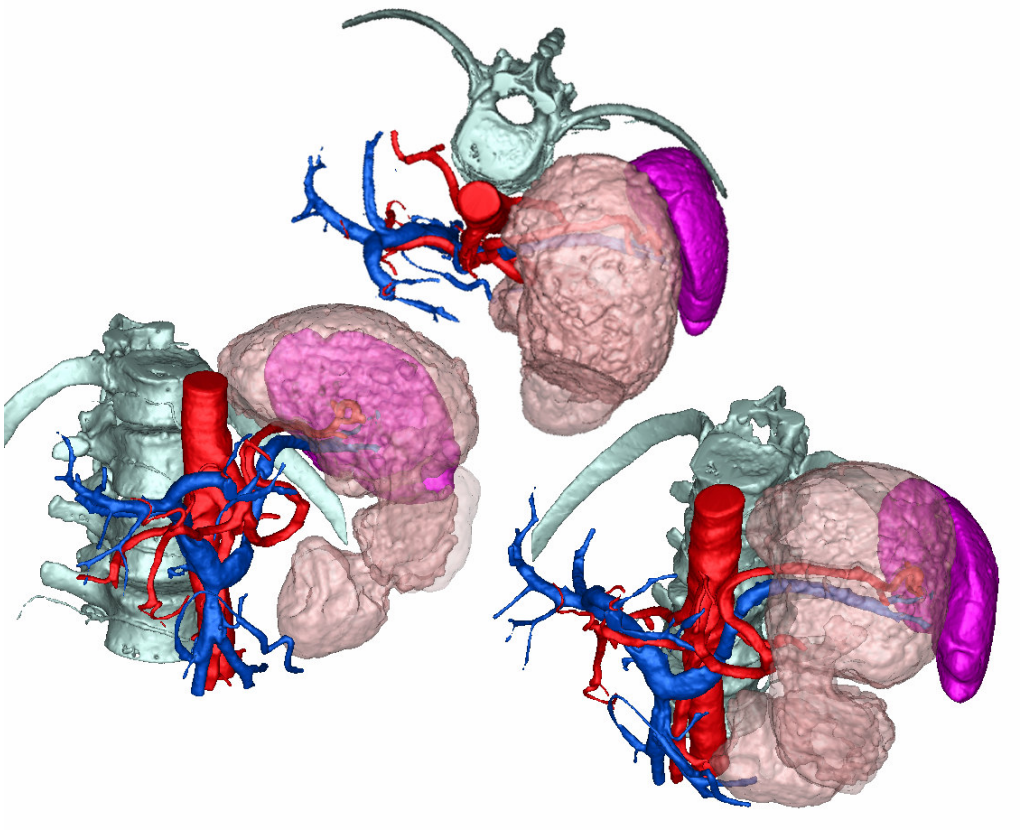

Fig 26 3D model reconstructed from real data acquired from a patient with the stomach full of water. The stomach (pink-transparent) touches the spleen (violet) and lies on splenic vessels. If the stomach would be empty, the spleen would be more inside and more up.

The beating of the heart introduces troubles not only in cardio surgery but also in vascular surgery, where the rhythmic changes of the blood pressure introduces visible deformations on arterial walls. The effect of the insufflations in laparoscopy can be considered a variation in the abdominal pressure. Its effect on the abdominal wall is reported on [52]. In general all physiological parameters (Young-Modul of each structures, blood pressure, muscle contractions, position and velocity of elements, etc…) can be a source of movement that would have to be taken into account.

Finally, another class of dissimilarities appear when, during the intervention, the surgeon modify the topology of the real anatomy. In this case the pre-operative information will not be completely consistent with the real anatomy and the virtual model would have to be adjusted. An example of adaptation of the virtual anatomy is shown in Fig. 22 (up). In this case the adaptation has been pre determinate by the surgeon (using virtual cutting plane). However this (useful and interesting) problem will not threaded and considerate in this thesis.

4.2.2 Decoupling the problem

On the base of the previous consideration and taking into account practical physicians needs we can try to formalize the problem grouping factors that influence the registration. The following considerations allow to write the problem in symbolic terms that can simplify the reading.

bc() is the bed configuration in terms of angles relative to the floor, shape and covering material, *dec()* the decubit of the patient on the bed and *pp()* all physiological parameters. The goal is the estimation in the surgical room of the trajectories of all points $\hat{V}sr_i(t)$, corresponding to the center of mass of each voxel Vrd_i (acquired in the pre-operative phase), in order to predict their real position:

$$\hat{V}sr_i(t) \cong Vsr_i\left(bc(t), dec(t), pp(t)\right) \qquad (1)$$

3D radiological devices acquire voxels "frozen" them in particular instants and often there are delays in the acquisition of the various voxels of an image. For example MDCT acquires some slices (depending on the number of detectors) at the same time, and after it acquires another portion of volume, and so on. In general we can consider that the barycentre of a generic voxel *i* is acquired at time t_i and depends on $bc(t_i)$, $dec(t_i)$ and $pp(t_i)$ of the patient at the acquisition time. If we do not care on artefacts introduced by the radiological device and eventually reconstruction algorithms, we can write:

$$Vrd_i \equiv Vrd_i\left(bp(t_i), dec(t_i), pp(t_i)\right) \qquad (2)$$

We have to estimate all $\hat{V}sr_i(t)$, using the pre-operative information contained in all Vrd_i (and eventually other sources of information), respecting *(1)* :

$$\forall \bar{t}, i \quad \hat{V}sr_i(\bar{t}) = F(Vrd_1 \ldots Vrd_n, \ldots) \mid \hat{V}sr_i(\bar{t}) \cong Vsr_i\left(bp(\bar{t}), dec(\bar{t}), pp(\bar{t})\right) \qquad (3)$$

From *(3)* we can write a weaker condition:

$$\forall \bar{t}, i, j \quad \left\| \hat{V}sr_i(\bar{t}) - \hat{V}sr_j(\bar{t}) \right\| \cong \left\| Vsr_i(\bar{t}) - Vsr_j(\bar{t}) \right\| \qquad (4)$$

that means that the configurations, in terms of distances, between real and estimated points as to be the similar as more as possible. Relation *(4)* remains valid applying rigid transformations. As consequence the problem can be in theory decomposed: for each instant we can calculate a rigid transformation, for the alignment of the two reference frame (*sr* and *rd*), followed by an adaptation of the pre-operative information to the real configuration (in

terms of relative distances like in *(4)*), to calculate instantaneously for each element of the patient.

This approach is used in every non rigid registration method [20] [48], independently by the motivations proposed here, because it is the more reasonable, with the only difference that they align two static cloud points, while in *(3)* a static point cloud, corresponding to the pre-operative information (Vrd_i points), has to be aligned with a "moving" point cloud corresponding to the intra-operative information (moving $Vsr_i(t)$ points).

The order of the two operations, rigid transformation and configuration adaptation, could be in theory exchanged, but, as follow motivated, the (more reasonable) classical order, used in non rigid registration, has to be preferred also for the solution of a dynamic non rigid registration.

It is obvious that we can not describe points trajectories $Vsr_i(t)$ using only the static values of all Vrd_i. Others intra-operative information are necessary. Many intra-operative information could be acquired using radiological volumetric devices (with eventually high refresh rate). This information would describe the trajectories of many points of the patient body during the intervention. The trajectories of the points not acquired by the intra-operative radiological device could be estimated (using interpolation, for example). But almost all surgical rooms has no volumetric radiological devices (like open MRI or similar). However the acquisition of some intra operative information, like that acquired by means a tracking system on the subject skin, can offer important conditions for biomechanical/physiological models similar to [43] [44] that demonstrate the possibility to predict point positions during breathing. Further we could build models that take into account others physiological parameters and eventually decubit of the patient and bed arrangement too. In other words measuring only part of the intra operative information, we could estimate the others modelling movements, deformations, etc. This estimation could be enhanced during the intervention, depending on the precision and on the type of intra-operative information available. For example, when the surgeon discover new decipherable anatomical landmarks, with localized instruments, with human eye, or better, directly by machine vision algorithms, this information could offer the position of same Vsr_i that could be used as useful new conditions for biomechanical or motion models. All this information will be represented in a reference frame of the surgical room. If we will develop a model, which will completely simulate a real abdomen in terms of mechanics relation (writing a differential equation system, following FEM methods, for example), it will be valid independently by the used reference frame. Otherwise using motion models like [53], for example, that simulate trajectories and not mechanics relations, it has to

be used the particular reference frame for which the trajectories have been defined. In the last case, problem decoupling, in rigid transformation and configuration adaptation, in this order, would be useful in order to obtain the desiderated reference frame.

The rigid transformation can be done aligning fiducial points [54] or fiducial surfaces [55] acquired in the radiological department and in the surgical room, considering the object as rigid. Deformations of the fiducial structure composed by elements like points of a cloud or points characterizing a surface, introduce systematic errors in the registration. In order to minimize the registration error, at least on fiducial elements, we have to choose each fiducial point (or fiducial surface) in the proximities of steady element on the patient and their configuration has to be replicable, as more as possible. In this case a rigid registration can be performed at the beginning of the intervention and should be repeated only if the bed configuration, or the patient decubit, will be changed.

A rigid transformation on Vrd_i can be written, using cartesian coordinates, in terms of rotation (matrix R) and translation (vector T):

$$\tilde{V}sr_i\left(bp(t_i),dec(t_i),pp(t_i)\right) = R \cdot Vrd_i\left(bp(t_i),dec(t_i),pp(t_i)\right) + T \qquad (5)$$

After that, the configuration adaptation becomes:

$$\forall \bar{t}, i \quad \hat{V}sr_i(\bar{t}) = F(\tilde{V}sr_i\left(bp(t_i),dec(t_i),pp(t_i)\right),\ldots) \qquad (6)$$

that has to verify *(4)*.

This last complex work, very difficult to solve entirely, can be facilitated with practical and useful artifices. These artifices are used every day by radiotherapists reproducing meticulously patient's setting during the treatment as in the planning room. Following their work, bed positioning and shape, during the acquisition of medical datasets, can be chosen accordingly to the bed configuration used inside the surgical room for the specific intervention (taking in mind the requirements of the radiological device used and the type of intervention to do). In this case the *bp* registration error source can be significantly reduced. Also replying exactly, during the intervention, the exact decubit of the patient during radiological scansion (in order to reduce *dec* error component), requires obtaining the same relative position of the basin and the thoracic case. A realignment of these structures needs additional iterative work in the surgical room in order to find a perfect correspondence between pre-operative and intra-operative patient positioning [56]. Further the position of arts influences the arrangement of abdominal organs and in particular: arms position influence especially upper abdomen structures, while legs position influence especially structures

situated in lower abdomen. In fact radiotherapists immobilize arms or legs depending on the zone that they has to tract.

Following all these considerations the model to build, for the simulation of patient's anatomy movements, will be drastically simplified, removing the *bp* ad *dec* component. Furthermore, like demonstrated by experimental results [57] (without patient moving), the application of models that predict organ shape motion, controlled by some parameters, like the time over the cardiac cycle, or the displacement of the diaphragm, can be applicable in the real surgical scenario with enough precision. A simple example of motion model is described in 4.4, while the focus of this thesis is the initial marker based rigid registration, described in the next paragraph.

4.3 RIGID MARKER BASED SOLUTION FOR THE INITIAL REGISTRATION

Following the previous considerations, it has been studied a usable initial rigid body registration method and it has been estimated the *dec* error component, in order to know the order of magnitude of inaccuracies when a perfect realignment of the basin with the thoracic case, using immobilization devices, is skipped. This experiment allowed also estimating the order of magnitude of the error due to breathing, on the patient abdominal skin, and to examine the methods to employ for the rigid registration of moving structures (not rigid).

4.3.1 Experiment for registration error estimation on the skin

In order to estimate the influence of the misalignment of the abdominal surrounding structures (in particular the basin and the thoracic case) on the registration error, it has been studied a real case. The trajectories of some markers, visible by an optical localizator, attached on the abdominal skin of a 30 years old male with weight 74 Kg and stature 182 cm, have been acquired and studied.

Performing the experiment with a slim subject allowed to use a simple marker based registration approach, since points on the skin were very steady especially in the proximity of promising bones. Otherwise we would have to use imaging devices, like US, to detect the reference internal rigid structures (bones) and proceed with a surface registration [58].

Basing on anatomical and technical considerations discussed by physicians and engineers, 18 wired infra red markers have been posed on specific anatomical points on the skin of the subject (Fig. 27). During the experiment, the "patient" has been repositioned many times on the surgical bed in supine position trying to reproduce roughly the same positioning between thorax and basin, correcting visible erroneous postures. Subject's markers have been acquired with an optic localizer (Optotrak Certus) at each repositioning for 1 minute. During the acquisition period the subject breathed normally. From the analysis of the acquired data it resulted that the abdominal surface of the subject has a maximum excursion, relative to the median point of each marker, of ±7.5 mm. This result agrees with [49]. Furthermore the 3 markers signed with a cross in Fig. 27 (the anterior superior iliac spine and the upper sternum) were resulted quite steady during breathing motion: 2.5 mm for the marker on upper sternum, while the others were very steady (< 1 mm). These results are due to anatomical positions of these markers that result marginally influenced by breathing motion.

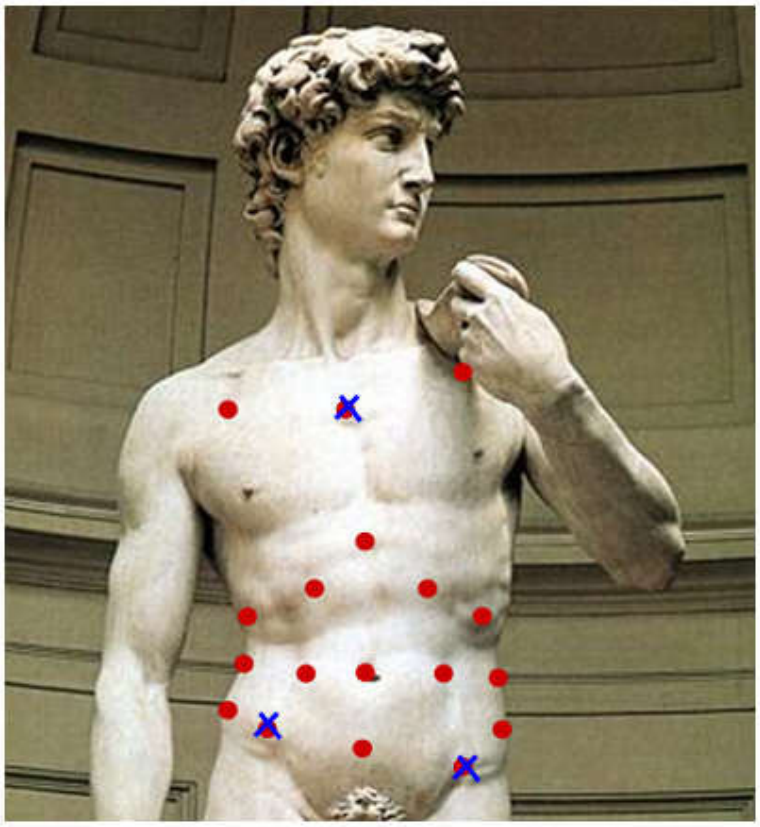

Fig. 27 Marker configuration on patient skin for movement evaluation (points), and the markers, more stable during breathing (crosses), chosen for registration.

Since registration errors prediction using an analytical approach, like in [59], cannot be adopted due to the non rigid nature of the abdomen, an heuristic approach has been used, using some points as fiducials (for the registration) and the others as potential targets (for error evaluation). In the next paragraphs results of this evaluation, which provide the order of magnitude of errors introduced when a perfect realigning procedure is not employed, are reported.

Since the application of CAS systems has to match with the real clinical scenario, which in general uses its procedures studied for "not computer aided surgery", markers have been also acquired in semi-lateral decubit. The evaluation of these data furnish the order of magnitude for the registration error if we try to apply CAS systems for interventions with semi lateral decubit (frequently used for abdominal interventions) where pre-operative images are acquired, without considering CAS needs, in the traditionally supine position.

4.3.2 Rigid-body registration algorithms

Using just a rigid approach we drastically approximate the solution, removing the *(6)* adaptation:

$$\hat{V}sr_i = R \cdot Vrd_i(bp(t_i), dec(t_i), pp(t_i)) + T \qquad (7)$$

It allows to adjust only the change of reference system: from the radiological device to the surgical room reference frame.

With this type of registration at each instant we have a registration error on each voxel *i*:

$$Ver_i(t) = \hat{V}sr_i - Vsr_i(t) = R \cdot Vrd_i(bp(t_i), dec(t_i), pp(t_i)) + T - Vsr_i(bc(t), dec(t), pp(t)) \qquad (8)$$

Now it is discussed the registration algorithm to use for the determination of R and T, necessary for the evaluation of registration errors on the patient skin, and in CAS applications, in general.

Matrix R and vector T of the equation *(7)* can be calculated, choosing in particular instants, values of the marker trajectories from the experiment and applying a marker based registration. In the next paragraphs will be shown how these two points cloud can be chosen. At the moment the fiducial points Frd_j can be considered as instantaneously acquired in the radiological department, and Fsr_j as instantaneously acquired in the surgical room.

R and T have to be chosen in order to:

$$\hat{F}sr_j = R \cdot Frd_j + T \Big| \quad \forall j \quad \hat{F}sr_j \cong Fsr_j \tag{9}$$

Fiducial localization error (FLE) errors on the localization of Frd_j and Fsr_j, and the deformation of structures to register (like in our case), determine that an optimal solution, that remap perfectly all Frd_j on the relative Fsr_j is impossible. We have a residual error on each fiducial:

$$Fer_j = Fsr_j - \hat{F}sr_j \tag{10}$$

so the solution of the registration problem is not unique, consequently we can not write the problem as an invertible system of equations. However we can find an optimal solution in function of the metric used for defining the goal of the problem.

If we consider each marker as unit of mass, we can align the baricenter and principal inertia axes of the two point cloud, using PCA (Principal Component Analysis) decomposition like in [49].

A method, based on least square fitting [54], is routinely used since the mathematical properties of the square norm, which allow the decomposition of the problem for finding translation and rotation [60]. Further it allows an analytic solution to the optimization problem, beyond that it is the optimal solution when FLE is gaussian. See the appendix A for a review of the least square solution using SVD decomposition, and appendix B for an exercise of the author that demonstrate, in case of no FLE, that SVD and PCA approaches offer (obviously) the same result.

Even if least square fitting offers an optimal solution in case of normal distributed FLE, it is deficient in case of non normal errors, in presence of outlier and when we have to control matching error on individual points (real cases for non rigid objects). For example using the 3 fiducial points we often have visible different punctual errors on the various markers and this is unexpected by the surgeon that would prefer to see, when the marker configuration allows it, a more natural equal distribution of the error when he moves the surgical instruments around fiducial markers (see last column of Tab. 2).

Frd_j: x,y,z (mm)	Fsr_j: x,y,z (mm)	Fer_j L2/Linf (mm)
-488.5, 74.6, -1549.6	-470.6, 2.1, -1578.4	2.7/2.9
-488.2, -153.0, -1557.2	-465.4, -220.8, -1589.9	3.5/2.9
-230.5, -44.8, -1895.0	-207.5, -112.9, -1924.0	1.1/2.9

Tab. 2 Registration error Fer_j from real data Frd_j and Fsr_j optimizing least square norm (*L2*) and infinite norm (*Linf*)

So registration matrixes have been calculated using least square fitting using SVD [54], in order to compare results with other works that commonly use this approach, and also using a sub optimal solution that optimize the error in terms of infinite norm [61], in order to obtain the same displacement on each fiducial and eventually to increase the margin of security during interventions.

4.3.3 Registration error estimation

Various possible methods to use for the accuracy evaluation, when we have to align deformable and moving structures (like in the experiment with markers on the skin of the patient) using a rigid approach, are now examined.

The estimation of the error introduced by different decubit, by means of the evaluation of data acquired with the optical localizer, has been performed registering data of each repositioning of the subject to data of an acquisition selected as reference (as it would be the patient in the radiological department). It requires to chose: markers to use as fiducials, particular values during their trajectories (see Fig. 29) to use as points for the registration algorithm, the method for calculating the registration matrix, and finally a way for quantifying the misalignment between registered moving data, of each repositioning, in respect to reference points (on the acquisition selected as reference). About the registration algorithm, R and T are computed, like described in 4.3.2, using norm-2 and norm-inf fitting. The others choices require various consideration as follow discussed.

4.3.3.1 Markers to use as fiducials

In a first step all markers have been considered fiducial markers, in order to estimate accurately the *dec* error component. We have to take into account that in the surgical room the acquisition of fiducial points could be difficult using active markers (Fig. 28 left) and a less invasive sterile digitalizer (like the one in Fig. 28 right) would be preferable by the medical staff because it do not introduce changes in the surgical scenario. As counter part the use of a digitizer requires additional work for the surgeon and generates errors in the acquisition of not steady points during breathing (especially on not slim patients), so the error considering only the 3 more stable points (corresponding to crosses on Fig. 27) has been also evaluated.

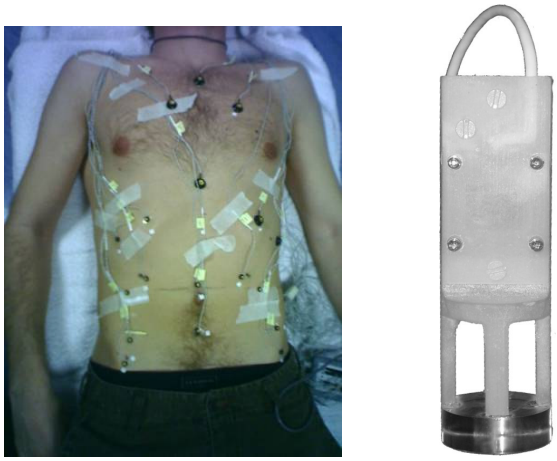

Fig. 28 (left) Wired markers on patient skin for the acquisition of trajectories of points. (right) Cutaneous digitizer for Optotrack Certus developed for the EndoCAS Navigator. The plexiglass base (with a reference on its surface) allows the acquisition reducing errors due to falling down in the skin like in standard "pen like" digitizer.

4.3.3.2 Values to use for the registration

The registration matrix, for the estimation of the error introduced by different decubit, can be calculated using as point cloud the instantaneous positions of the fiducial markers in any instant of the breathing cycle, or in their mean position (like in [62]).

Since values of the experiment can be used for error evaluation due to breathing, we can choose values of the trajectories in two significant ways. The first one takes into account real conditions in the radiological scenario, where acquisitions, using breathing trigger, are done in the exhalation phase (due to the relative stability of this point during the breathing cycle). For this reason the instant times, were the marker on the umbilicus reach the minimum distance from the bed, have been calculated (Fig. 29). After that, the mean of positions during these instants (of each fiducial marker) has been assumed as fiducial point Frd_j, for the acquisition chosen as reference, and as Fsr_j, for the others repositioning of the subject.

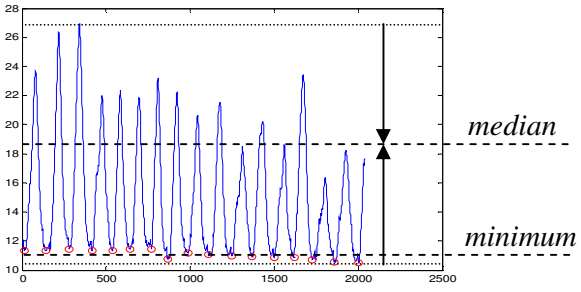

Fig. 29 Vertical displacement of the marker on the umbilicus and the two values used for the registration.

In the second case it has been considered the median point, the middle of the extremely values reached by each marker. This approach intend to minimize the registration error during breathing.

4.3.3.3 Various usable registration error definitions

Using the previous consideration many R and T couples, to use as in *(7)*, have been calculated.

How shown in *(8)* the registration error on targets depends on the value assumed as acquisition time t_i for each target voxel Vrd_i, in the radiological department, and on the movement of the targets $Vsr_i(t)$ during the intervention.

All 18 acquired markers are considered potential targets, and various error definitions are used as following defined:

mean: the error is calculated directly on the 18 points Fri_i and Fsr_i, used for the registration, without considering specific values of the trajectories. For each patient repositioning the least square error (TRE in [59]) was calculated and followed by a mean on all results. This error definition is a sort of mean for all target errors, where the breathing component is compensated.

maximum: is calculated directly on the 18 points Fri_i and Fsr_i, used for the registration, without considering specific values of the trajectories. For each patient repositioning it was calculated the maximum of all Fer_j defined in *(10)*. Finally it was calculated the maximum on all partial values. This error definition can be considered an upper limit for the registration error of points that are not influenced by breathing.

Dynamic: is calculated as the maximum value of the trajectory error $Ver_i(t)$ (defined in *(8)*) of all targets and for all repositioning of the patient of the same type (supine or semi-lateral decubit). This error definition provide an upper limit for inaccuracies in presence of breathing.

4.3.4 Rigid registration evaluation

Using various approaches, like previously described, registration errors of the subject, repositioned many times on the surgical bed in supine position, have been calculated. The results are shown in Tab. 3.

Fiducial Markers	Fiducial Values	Mean error L2/Linf (mm)	Maximum error L2/Linf (mm)	Dynamic error L2/Linf (mm)
All	*median*	*2.3/2.6*	*5.1/4.2*	*9.7/10.8*
3 MS	*median*	*3.0/3.1*	*6.3/6.1*	*11.3/11.3*
All	*minimun*	*2.2/2.5*	*5.1/4.1*	*16.9/17.6*
3 MS	*minimun*	*2.9/3.1*	*6.1/6.1*	*17.9/17.9*

Tab. 3 Registration error obtained without perfect realignment, correcting only visible misalignments, on a slim breathing subject in supine position optimizing square norm (*L2*) and infinite norm (*Linf*)

At first we can see that using all markers as fiducials we obtain a *maximum* error <=5.1 mm, that can be considered an upper limit for the registration error of steady points during breathing. Although when we do not control directly the errors on all targets, using only the 3 most stable markers, this error increases just up to 6.3 mm. Also results of *mean* error (that offer a sort of mean error on steady structures) sustain the use of the 3 most stable markers as fiducial, obtaining in this case an accuracy very close to the one found using all markers. In both cases *mean* and *maximum* error, compensating breathing effect, provide the order of magnitude of the misalignment due to not perfect realignment of the basin with the thoracic case. We can see that this source of error results not too great, so in many CAS applications (depending on the accuracy required) the perfect patient realignment, using immobilization devices, can be skipped.

Obviously, reproducing the same experiment, with more fat patients, the error will result higher due to movements of fiducial markers. But, presumably, the error caused by the not prefect alignment between basin and thorax will be approximately similar. So, using intra operative imaging for finding internal steady markers or surfaces, we can expect that the error will have the same order of magnitude on all patients (fat or thin).

Analyzing *dynamic* error, we can see that using the median values of fiducial trajectories for *R* and *T* calculation, the dynamic error could be significantly reduced as supposed (respect to the use of minimum values). But this result can be considered (today) only a theoretical optimal goal. In fact radiological devices offered by the market have at least breathing trigger in the exhalation phase (there are some technical and practical problems to solve in order to offer triggers in other instants). How shown the registration error depends on the acquisition instant chosen for the acquisition of the radiological dataset. This result has to be taken especially into account using radiological scanners not provided of any breathing trigger

device. Like shown in 4.2.1 using the inhalation apnea we can introduce an error up to 6-8 cm.

If we have to apply CAS aids for interventions with semi lateral decubit, where pre-operative images have been acquired in the traditionally supine position, we will have errors with an order of magnitude like shown in Tab. 4.

Fiducial Markers	Fiducial Values	Mean error L2/Linf (mm)	Maximum error L2/Linf (mm)	Dynamic error L2/Linf (mm)
3 MS	*minimun*	*17.3/16.5*	*39.4/37.7*	*40.1/38.5*
All	*minimun*	*14.6/14.9*	*24.7/21.6*	*25.4/22.6*

Tab. 4 Registration error evaluation between points acquired on the subject in semi-lateral decubit registered to points acquired in a reference supine position.

In this case, even optimizing the maximum error (using infinite norm), we will obtain accuracies probably not enough for computer aided surgery systems.

Regarding the use of the registration algorithm that optimizes the infinite norm, we can see in Tab. 3 that its accuracy (without great deformations) is comparable to the one obtained using least square norm. However infinite norm algorithm can offer more natural results for the surgeon, which will see the same error on all fiducial markers (note that if the deformation of fiducial configuration is too big, the algorithm, trying to reduce the error on points that have the maximum displacement, can not offer the same error on each marker).

4.3.5 A rigid registration procedure compatible with the clinical scenario

On the base of these results, a practical registration procedure, for not obese patients, can be:

1. application of visible spherical markers (radio opaque for CT) on the sternum, as high as possible (where the breathing component is lower), and on anterior superior iliac spines;
2. acquisition of the preoperative dataset reducing all possible sources of errors, and calculating all fiducial points Frd_i as marker baricenter.
3. patient positioning, in the surgical room, as more as possible like in the radiological department, and acquisition (using a digitizer like the one in Fig. 28 right) of the points Fsr_i corresponding to markers position;

4. registration with square or infinite norm algorithm. EndoCAS navigator integrates both.

4.3.6 Application of the rigid registration procedure on a real case

The procedure has been tested on a real case [12] to perform a laparoscopic distal pancreatectomy. The maximum registration error on the 3 fiducials was 7.3 mm, which was 3.3 mm in the experiment using active marker attached on the skin. This increment is consequence of the FLE obtained in the real clinical scenario, using radio opaque spherical markers (in the radiological department) and the digitizer (in the surgical room), instead of active markers on the skin like in the ideal scenario of the experiment. This consideration has to be considered in the evaluation of the global result. The registration error on the target (a tumor in the pancreas) was estimated in about 2 cm. It is important to notice that the pancreas is a quite floating anatomical structure so 2 cm can be assumed as an estimation of the worst case error. Taking into account the error on fiducials and that the physiological movement can not be controlled, this result is coherent with results of the experiment with the optical localizer.

In conclusion registration of the abdomen using a rigid approach, even if it is not adequate for automatic robotic intervention, where a high level of accuracy is required, is considered adequate by surgeons for the development of intra-operative navigation systems. Furthermore it can offer a good starting point for more accurate deformable approaches.

4.3.7 Error estimation inside the abdomen

The rigid registration method described in 4.3.4 has been proposed on the base of data acquired on the patient's skin and having only few internal data from the intervention performed.

Like introduced previously, US is an optimal methods for acquiring internal points. It allows to acquire intra-operative and intra-body information with a simple (and cheap) device. Further the continuously advances of US devices increase every day the possibilities of its using.

The acquisition of 3D information using 2D US requires a tracking system and a sensor (calibrated with the US image plane) mounted on the probe. Using the US module of

EndoCAS Navigator, comprising tracker interface and routines for probe calibration, some experiments have been performed in order to estimate the registration error on internal targets using the rigid methods described in 4.3.5.

At first 3 subjects has been scanned in order to find points to use as anatomical internal targets (in the exhalation phase). In all cases a target is chosen in correspondence of the bifurcation of the iliac arteries (lower abdomen). This repere can be acquired position the probe in coronal position and finding the plane where the circular sections of the iliac arteries touching together on the abdominal aorta. The target points have been manually selected on the US plane. Others target points have been chosen, one in the spleen, and one in the liver (upper abdomen), on the base of the anatomy of each subject (points simple to recognize).

After that patient's skin has been signed on the 3 points used for the registration and the patient has been re-positioned and acquired many times on a surgical bed. The points on the skin, acquired with the digitizer, have been used for the registration, while the chosen internal points, acquired with the US, have been used as target points for the registration error estimation.

In all cases the error has been lower than 8 mm. This result, considering the difficulties in US localization (image noise and pressure on the probe to find the target), validate the rigid registration error obtained on the skin in the exhalation phase (see cell 4-4 on Tab. 3).

It offer also an important accuracy estimation method for CAS applications.

4.4 AN EXAMPLE OF MOVING MODEL: PULSING ARTERY

As descript in 4.2.2 motion models can be useful and usable for movements simulation for registration purposes. A simple approach to the problem can be performed modifying the primitives of the surfaces generated from the radiological dataset. For demonstration an application that simulates the pulsation movement on abdominal arteries surface, has been developed on the top of EndoCAS Navigator platform. The movement is simulated moving the vertex of the triangular mesh (generated as report in the previous chapter), as done in many applications, for example in talking head simulation [63] [64]. For the pulsing effect simulation, each vertex is moved along its normal direction (automatically calculated by the OpenSG routines used in the navigator), by a value function of the time (using a function composed with positive parts of sinusoids, that empirically simulate the pulsation effect).

Even if the solution proposed as no physical relationship, the result has been appreciated by surgeons that seen a first verisimilar movement based on real data. The pulse simulation significantly enhance the visualization respect to the red coloured meshes commonly used for the arteries. This simple application can also be interfaced with a pulsimeter, in order to offer a pulsation synchronous with the patient heart signal. In this way the surgeon will increase the perception of model correspondence with the reality, and probably he/her will take more attention when he/her move instruments close arteries.

5.0 ENHANCING CAS POTENTIALITIES ELABORATING CAMERA IMAGES

The work reported in this chapter derives from some considerations of the previous one, regarding the registration of the abdomen. The analysis of the registration problem show that working without the use of volumetric intra-operative imaging, we encounter experienced a lack of information for the modelling of all abdominal deformations, necessary for a complete non rigid registration task. A large amount of internal intra-operative information can be acquired with the elaboration of laparoscopic images. Furthermore, this information could be useful for the refinement of the initial rigid registration, which is especially useful for obese patients, where marker mobility can compromise the alignment accuracy.

For these reasons, monoscopic and stereoscopic 3D information extraction approaches, using laparoscope and cameras in general, have been developed. They also allowed to introduce new potentialities for the CAS scenario, using mixed-reality, described in this chapter.

5.1 BACKGROUND

All machine vision algorithms that elaborate 3D information require to know the type of camera used and its internal model in terms of ray trajectories, from the world to the camera sensor. Line scan and telecentric camera types are used for particular industrial applications, while for all visualization purposes, including laparoscopy, the perspective projective camera is the only used, because it offers the most similar images in respect to human vision.

Regarding the sensor, two technologies are predominantly applied: CCD (Charge Coupled Device) and CMOS (Complementary Metal Oxide Semiconductor). In each case unitary elements (pixels) are disposed on a regular grid (with fixed resolution). These technologies differ only in the acquisition frame rate and the quality which can be obtained in case of moving objects. Therefore, they do not influence the parameterization of the internal model, which depends just on the pixels arrangement and not on the pixel technology.

The following dissertation is independent of the perspective cameras used, which can be modeled as pinhole, as shown in the next paragraph.

5.1.1 The pinhole camera model

Each camera, composed of a projective optics and a grid sensor, can be represented by the following model:

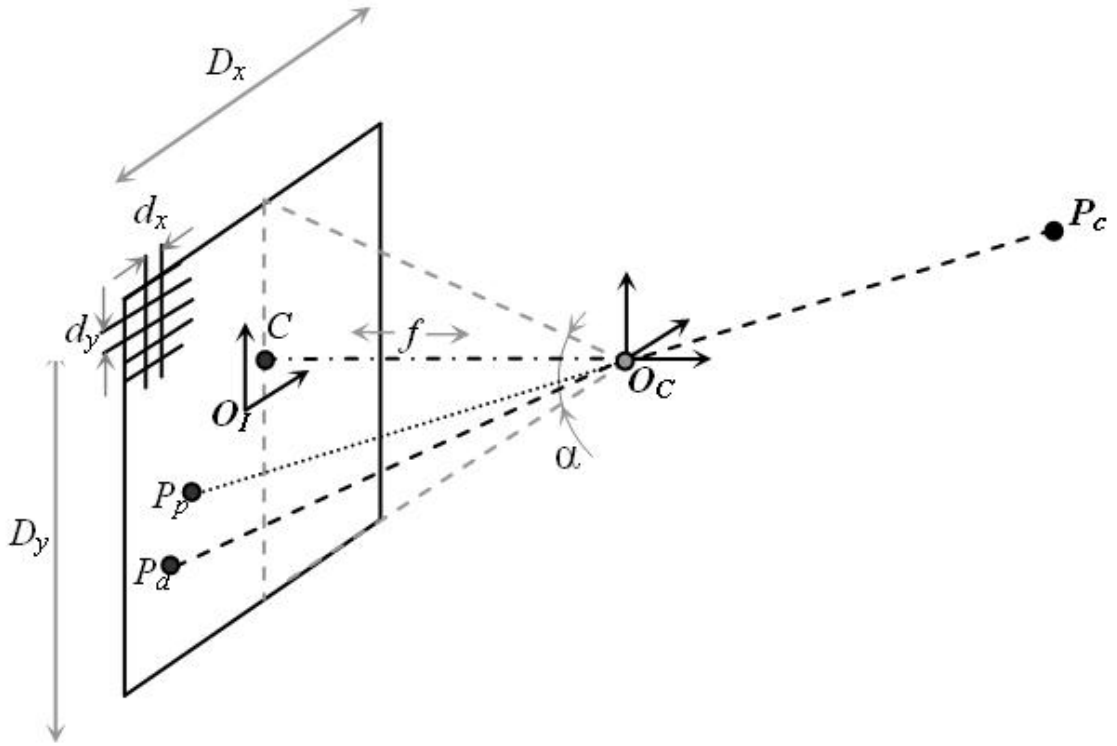

Fig. 30 Schematic representation of the pinhole camera model: the generic point *Pc* is ideally projected on the image sensor of the camera (the plane with origin O_I) through the projection center O_C (where the origin of the camera reference frame is fixed).

The perspective projection matrix Mp, mapping a generic 3D point Pc = [x, y, z, 1]^T, in the camera reference system, to the corresponding 2D point Pp = [u, v, 1]^T in the image reference system (fixed on the center of the CCD), i.e.:

$$P_p = M_p P_c \tag{11}$$

is defined starting from the internal camera parameters (f, C_x, C_y) as follows:

$$M_p = \begin{bmatrix} -f & 0 & Cx & 0 \\ 0 & -f & Cy & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{12}$$

where f is the focal distance and (C_x, C_y) are the coordinates of the projection of the O_c on the image reference frame (with origin in O_I).

The pixelization process is defined by the pixel dimensions d_x and d_y and the image sensor dimensions D_x and D_y. These internal parameters of the camera allow to convert measurements done on the image (in pixels) in real measurements (in millimeters) and vice-versa.

The last internal camera parameters parameterize the model of the radial distortion, introduced by common lens, by means of which the projected point P_p is deviated on P_d.

All internal camera parameters can be determined in a calibration phase acquiring some images of a knowing object in different positions with fixed camera configuration (in terms of diaphragm and camera focus) and using calibration routines like [14]. A new calibration, which takes just few minutes, is required whenever either the optic zoom or the diaphragm opening is changed.

5.1.2 The mixed-reality concept

Mixed-reality allows the user to see "augmented images", crated mixing virtual pre-operative information obtained processing radiological images (CT or MRI) with real patient live images, like shown in Fig. 31, real or grabbed by means of cameras, in a natural way [65].

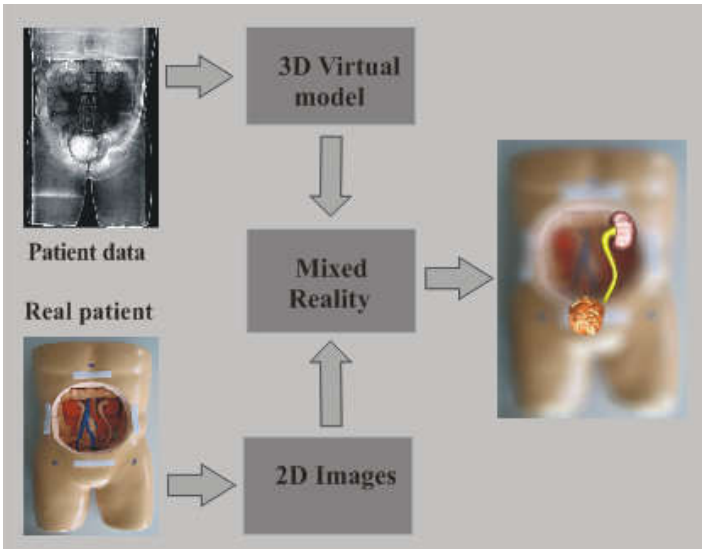

Fig. 31 The mixed-reality concept.

Since preoperative information, in the form of 3D models, are virtually projected onto real images in the corresponding position, the user has the sensation to see-through the patient. It

introduces many advantages for each task where the physician have to interact with the patient (palpation, introduction of needles or catheters, intervention, etc.).

In order to mix coherently the virtual information with the real information, each virtual camera (part of the rendering mechanism) has to be modelled identical to the corresponding real camera, and the virtual images has to be aligned to the real ones (Fig. 32).

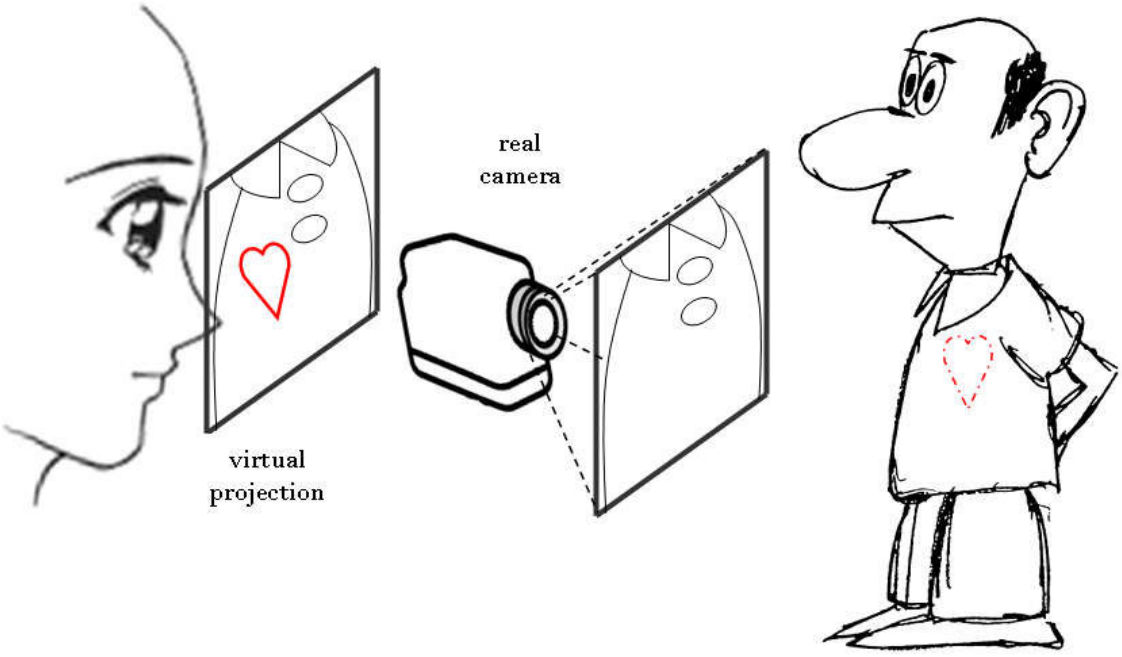

Fig. 32 Mixed-reality and see-through. Real images, grabbed by means of a camera, are added with additional virtual information. If the virtual information is obtained from the real world (in our case using a radiological scanner), the virtual projection offers a see-through visualization.

The mixing of the real images and the virtual rendering can be done using hardware video mixer or using the real images in the scene graph as foreground or background [12].

Monoscopic mixed-reality enhances physician understanding but has some visual limitations especially regarding depth perception [66]. Depth perception can be drastically increased using head mounted stereoscopic see-through devices [67] that allow to appreciate object depth dislocation, like in the natural binocular view. HMD has also the advantage to allow the physician to see the scene in a natural way, with a point of view aligned with eyes, and, through helmet localization, offers the possibility to change the point of view just moving the head. For mixed-reality implementation, the video see-through approach, based on the acquisition of real images by means of external cameras, is preferable to the optic see-through approach that projects virtual information on semi transparent glasses. This is due to the fact that tracking of eye movements (which humans use for small gaze changing or to

anticipate head movement), strictly required for optical see-through approach, is very difficult to be performed with sufficient precision [68] [69]. On the contrary, head tracking, required for video see-through approach, can be performed with high precision using external localizer based on different technologies [70] [71].

5.1.3 3D localization using cameras

3D localization using cameras is used for many applications. Several computer vision libraries[72] offers many tools for this purpose.

Using a single camera, we can localize objects with known geometry [73] as in the case of EasyOn by Seac02 *(www.seac02.it)*. The localization accuracy is enough for many applications, but requires knowing in advance the dimensions and the texture of a rigid object in the scene. Interesting monoscopic solutions have been applied using laparoscopic images: localizing organs using artificial markers [74], recovering the position of needle [75] and the pose of surgical instruments [76].

Epipolar geometry [77], using two or more cameras, allows to detect the 3D position of each conjugate points, identifiable in the images. In a stereoscopic configuration, knowing the internal camera parameters, for each marker position P_d, in the image plane, the relative projection line in the 3D world, defined as the line l passing through the camera center of projection O_c and lying on the point P_c, is determined. These steps, performed both on left and right images, identify respectively two projection lines l_l and l_r. Knowing the relative pose of the right camera to the left camera (expressed by a roto-traslation matrix M_{r2l}), the 3D position of each marker is then defined as the intersection point between l_l and l_r. Since l_l and l_r do not intersect (due to pixelization process and noise in marker identification) the 3D marker position is approximated with the position of the closest point to both projection lines.

For stereoscopic localization, the external parameters contained in the matrix M_{r2l} that can be determined acquiring a knowing object with fixed cameras configuration (in terms of vergence and in terms of optic camera settings) are required, in addition to the internal parameters of each camera.

An interesting localization approach is proposed by [78] localizing the laparoscope. In this case the stereoscopic localization is performed using a single (moving) camera. A localizer offers the displacement of the point of view among the different images. If the surgical scene

does not change quickly (as in a real scenario) one has all ingredients for stereoscopic localization.

5.2 MIXED-REALITY USING LAPAROSCOPE AND HMD

5.2.1 System setup

Monoscopic and stereoscopic mixed-reality views have been developed and integrated in the EndoCAS Navigator system. The virtual objects used in the additional views are generated from real radiological data as described in Chapter 3.

The monoscopic implementation has been developed using the laparoscope as image source (coupled with a frame grabber) and a monitor for the visualization. The stereoscopic system is implemented using a custom made stereoscopic video see-through head mounted display (HMD) (Fig. 33). The helmet comprises an HMD (nVisor SX by NVIS Inc., www.nvisinc.com) with two internal SXGA LCD monitors, and two colour USB SXGA cameras (IDS uEye UI-1646LE) with 1/3'' CCD and 9 mm focal length optic. Focal length is chosen to respect the distances perceptual relations. Cameras are attached to the HMD by means of a mechanical support that allows to adjust the cameras vergence in order to assure stereo perception at every distance from the focused object.

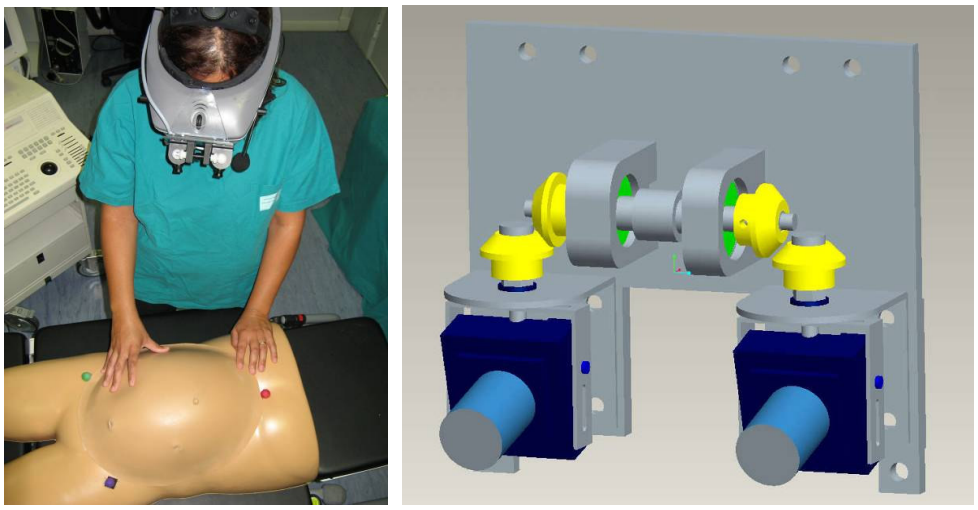

Fig. 33 The mechanical support mounting the cameras: prototype (left) and CAD drawings (right).

As written previously, in order to implement mixed-reality we need to provide a virtual model of the real cameras and its movements. In each case (laparoscope and HMD), the cameras are localized using the Optotrack Certus tracker integrated in the system, by means of frames equipped with infrared led, and the real patient is registered acquiring fiducial points on the skin using a digitizer as described in 4.3.4.

The mixed-reality mechanism, developed on the top of the EndoCAS Navigator, uses the software framework OpenSG 1.8 *(www.opensg.org)* for scene graph implementation and management. In the scene graph, live images, grabbed by means of real cameras, are used as background in the viewports that render the virtual scene.

The mixed-reality mechanism is implemented projecting coherently the virtual objects on the viewports and, hence, superimposing virtual images on the live images of the background. The mixed-reality mechanism requires the definition of a virtual camera model that exactly reproduces the real one, and the alignment of live and virtually reconstructed images by means of patient and camera localization.

In the next paragraphs the modeling of the virtual cameras and the camera calibration method used for the sensorized frames are discussed.

5.2.2 Virtual cameras modeling

The real internal camera parameters, determined with a calibration process offered by Halcon libraries, have been used to model virtual cameras. The virtual cameras have been modelled as off-center perspective cameras using scene graph libraries OpenSG.

The technical implementation requires the definition of the vertical field of view angle α (see Fig. 30). Starting from the focal length f and the projection (C_x, C_y) of the camera reference frame origin O_c on the image plane, the angle α is determined by the equation:

$$\alpha = \arctan\left(\frac{C_y}{f}\right) + \arctan\left(\frac{D_y - C_y}{f}\right) \qquad (13)$$

The horizontal field of view is set automatically on the base of the sensor resolution and dimensions.

The implementation of the off-center perspective camera in OpenSG requires the change of the reference system and normalization of the projection of the O_c on the image reference frame. This is done by means of the equations:

$$C_x^{'} = \frac{2C_x - D_x}{D_x} \qquad C_y^{'} = -\frac{2C_y - D_y}{D_y} \tag{14}$$

The misalignment between real and virtual projection due to radial distortion have been removed compensating the deformation on the real camera images. This step is necessary to obtain exact alignment and realistic mixed-reality representation especially with cameras having great distortions (as in our case). Radial distortion correction has been implemented with OpenSG, adjusting directly the background of each camera. The original version of OpenSG (in the 1.8 release) does not allows to correct images directly, but it simulates radial distortion mapping the image as texture on a triangulated background plate, where triangle vertexes are moved respect to a regular grid. Regular grid deformation function, implemented in OpenSG for radial distortion, is the polynomial function used in [79]. The radial distortion model used by Halcon is totally different from the OpenSG model, and is given by the equation:

$$P_d = \frac{2}{1 + \sqrt{1 - 4k\left\|\overline{P_p O_i}\right\|^2}}\left(P_p - O_i\right) + O_i \tag{15}$$

where k is the radial distortion factor. Starting from *(15)* it is possible to recover analytically the non distorted P_p point for each distorted P_d point:

$$P_p = \frac{1}{1 + k\left\|\overline{P_d O_i}\right\|^2}\left(P_d - O_i\right) + O_i \tag{16}$$

The equation *(16)* can not be analytically translated in a polynomial function, so it is not compatible with the model implemented in OpenSG. So, on the top of OpenSG structures, I implemented a function for grid deformation based on the equation *(16)*, obtaining the correction of the radial distortion of the live images grabbed by cameras. This solution is very simple and fast since OpenSG uses GPU (graphical processing unit) to perform computations.

5.2.3 Frame/camera calibration

To reproduce coherently the real movements of the cameras in the virtual environment, a calibration method, that solves the geometrical problem described in Fig. 34, has been implemented. The calibration matrix (*Tc*), representing the relative position of the camera viewpoint with respect to the sensorized frame, has been computed using a calibration grid

sensorized with infrared leds. The transformations *T1* and *T2* are given by the localization system, while the transformation *T3* is determined using a computer vision method [72] that allows to localize, in the camera reference frame, objects with known geometry (the sensorized calibration grid).

The alignment error between the real image and the virtual image has been evaluated positioning the calibration grid (160 mm x 160 mm) perpendicular to the laparoscope point of view and at a distance of 150 mm. The camera point of view has been determined using the *T1Tc* transformation while *T3* has been determined using the infrared leds sensors on the grid. Under this experimental condition the maximum displacement between the 4 corners of the grid in the real image and the correspondent points on the superimposed virtual image (see Fig. 34 right) has been estimated in about 2 mm.

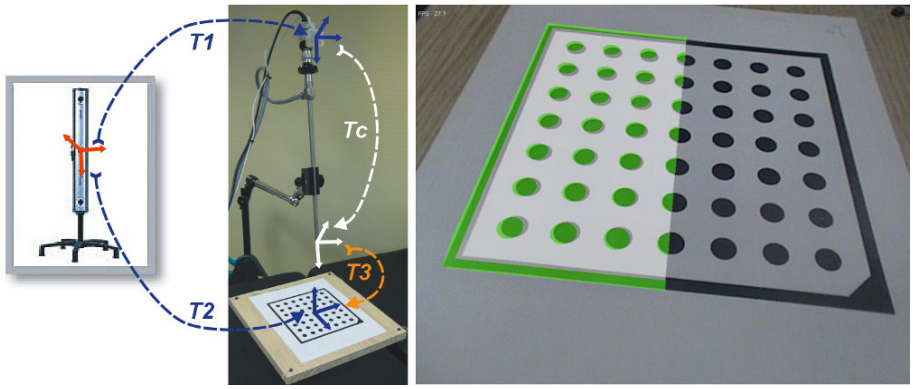

Fig. 34 Camera calibration process: (left) the calibration matrix Tc is calculated by solving the equation matrix represented graphically in the figure, where T1 and T2 are acquired with the localizer and T3 is computed acquiring with the camera an object with a predetermined reference frame and known features positions (the IRED sensorized calibration grid); (right) calibration result: mixed-reality laparoscopic view of the calibration grid (semitransparent virtual grid in green).

5.3 TRACKER-FREE STEREOSCOPIC VIDEO SEE-THROUGH

In the mixed-reality views, described in the previous paragraph, the camera localization is performed by means of an external tracker. As reported in the introduction, an external localization systems offers high precision, but, on the other hand, their use introduces a lot of problems concerning system setup, large footprint and needs of frequent calibrations. Further, the high cost of commercial trackers limits the diffusion of mixed-reality technologies on a

large scale.

For these reasons a tracker free version of the stereoscopic video see-through has been developed. The movements of the head of the user and the alignment of virtual patient are done using machine vision methods applied on pairs of live images (Fig. 35).

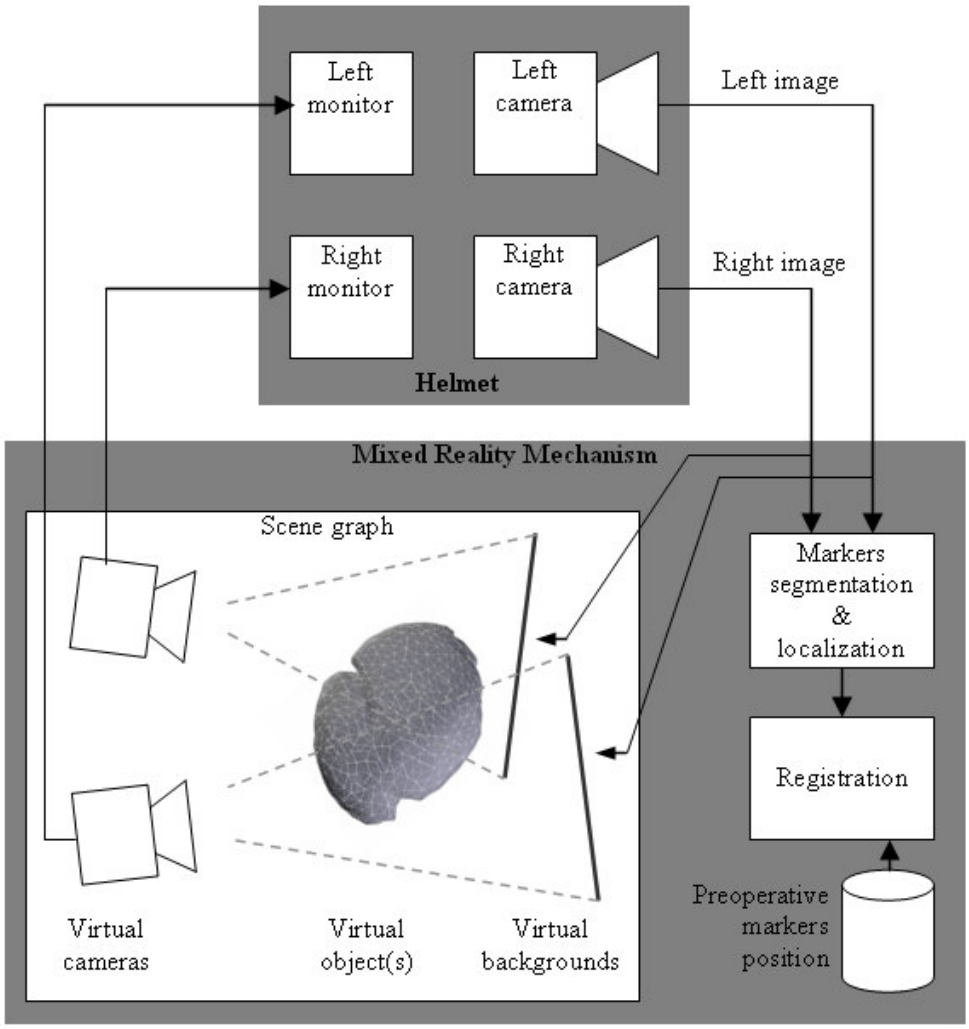

Fig. 35 Tracker-free version of mixed-reality HMD

Also in this case, 3 or more radio opaque markers have to be attached on the patient's skin before the acquisition of the medical dataset (MDCT), and their position (pre-operative fiducial markers position: $F_1 \ldots F_n$) has to be identified on the medical images during the segmentation phase. While real-time segmentation and localization of intra-operative fiducial coloured markers on the live images are performed by means of machine stereoscopic vision routines using the Halcon 7.1 software library developed by MVTech (www.mvtech.com). The alignment is obtained by applying a rigid registration [54] for each new intra-operative markers position onto pre-operative fiducial markers position.

5.3.1 Fiducial markers localization and registration

The alignment (superimposition) of the virtual world with the real world is made updating in real-time the roto-translation matrices M_l and M_r that describe the geometrical relation between the cameras reference frame (i.e. the helmet reference frame) and the scene graph reference frame (i.e. the medical dataset reference frame) (see Fig. 36).

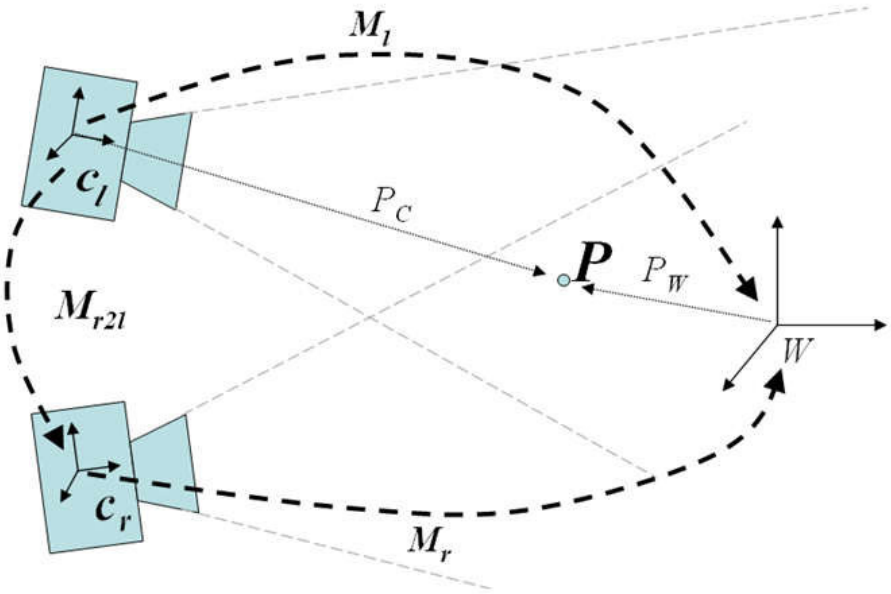

Fig. 36 Geometric transformation involved in the helmet.

Matrices M_l and M_r are defined registering the position of fiducial markers segmented on the medical dataset, determined preoperatively and given in medical dataset reference frame, to their real position, measured in real time using stereoscopic vision routines on pair of live images (left and right camera image). Artificial markers consisting in about 1 cm diameter felt balls coloured with different uniform colours (red, green and blue) are used. Felt material allows to sensibly reducing reflection artefacts, a frequent trouble for machine vision.

The real-time tracking of markers involves a segmentation process, performed on the grabbed images, consisting in two steps: colour segmentation and circular shape recognition.

Colour segmentation is performed using the HSV representation, in order to reduce the segmentation errors due to the artefacts introduced by the change of the illumination level, using a thresholding on the HS values for each marker [80]. Preset H and S thresholds values for each colour are defined after a training phase, and can be interactively adjusted. Transformation of the images from the RGB camera colour space to the HSV space is performed using a look-up table in order to speed up the frame rate [81]. After thresholding,

the circularity shape factor *CSF* is computed for each connected region of the colour segmented image. The circularity shape factor is defined as follow:

$$CSF = \frac{A}{D_{\max}{}^2 \pi} \tag{17}$$

where A is the area of the region and D_{max} the maximum distance from pixels to region barycenter (note that for a circular connected region $CSF = 1$).

Fiducial marker is then chosen as the biggest region having an $CSF > 0.5$. This empirical method identifies fiducial markers on the image plane with very low computational time. It was evaluated that the target identification fails in less than 2% of cases. This results is sufficient for our purposes since the registration is computed a lot of time in a second. Obviously the method definitely fails if large circular objects, having the same colour of markers is inserted in the scene, but this situation can be simply avoided in a real clinical scenario.

After marker segmentation, on the image plane, fiducial markers are localized in the 3D space applying stereoscopic vision routines. First of all the marker position is identified on the image plane with a single pixel consisting in the barycenter of the corresponding segmented region. Like descript in 5.1.3, knowing the internal camera parameters and the relative pose of the right camera to the left camera (M_{r2l}), the 3D position of each marker can be found, in the reference frame of the left camera (for example) obtaining the points F_{Cli}. Having the two points clouds composed respectively by fiducial markers in the radiological scanner reference frame F_i and fiducial markers in the reference frame of the left camera F_{Cli}, M_l is chosen in order to align as good as possible each couple of fiducials:

$$M_l F_i \cong F_{Cli}, \forall i \tag{18}$$

using the established registration algorithm [54], and M_r is obtained applying the fixed geometric relation M_{r2l} between cameras:

$$M_r = M_{r2l} M_l \tag{19}$$

The evaluation of the fiducial registration error (FRE) allows to individuate fault conditions in fiducial markers segmentation. In fact high values of FRE can be due just to an error in the stereoscopic localization of markers.

To reduce computational time required for marker segmentation, and to obtain 25 *fps*, the new marker position in the images can be searched in a sub-image square region centered in previous marker position and having an area proportional to FRE. When a localization error

occurs (a too large FRE) the entire image has to be processed one more time.

5.3.2 Performance evaluation

System evaluation was performed on a trunk phantom for laparoscopy by Sawbones *(www.sawbones.com)*. Radio opaque fiducial markers have been attached on the phantom: one on the sternum and two on positions corresponding to iliac spines. The phantom has been scanned with CT and surfaces of fiducial markers (for registration) and of anatomical structures (for mixed-reality visualization) have been extracted. Barycenters of segmented markers are used as fiducials F_i, while the corresponding F_{Cli} are obtained localizing with the stereoscopic routines felt balls attached in the corresponding positions.

The quality of the superimposition of the virtual anatomies to the real scene depends from the accuracy in localizing fiducial markers on the cameras' images sensors and from the cameras configuration. Error in localizing markers on image sensor has been estimated in $\Delta p = \pm 3$ pixel. The major component of the localization error due to camera configuration is measured along the z axis, and is defined analytically by the distance resolution formula [72]:

$$\Delta z = \frac{z^2}{f \cdot b} \Delta d \qquad (20)$$

where b is the distance between cameras, z represents the working distance, f is the focal length and Δd is given by Δp*$pixel\ dimension$.

In the system, where b=70 mm (anthropometric value), f = 9 mm, Δd=±11 μm, one obtains a Δz = ±4.4mm at a working distance z=500 mm. This error is principally due to the short distance between the cameras respect to the working distance. Even if the distance resolution error is numerically significant, in mixed-reality visualization it is perceived as a negligible alignment error from the user. Further in our lab, we experienced that the use of more precise localization systems does not introduce significant visualization improvement that are perceived as substantial from the user.

Results of the measurements and a consideration for the evaluation of the tracker-free system compared with using commercial localizers are reported in the following paragraphs.

In 5.2.3 it is shown that we can obtain very precise alignment localizing a camera with a very precise localizer (Optotrack Certus by NDI, the gold standard in commercial localizers) and

sensorizing both the camera and the real object (a calibrated grid representing conceptually the patient). The precision obtained under these ideal conditions significantly decreases in a real clinical scenario for two main reasons. First of all, the patient is not rigid and secondly localizable fiducial markers cannot be fixed steadily on the patient. The last condition requires human actions for the re-positioning of localizable markers on the patient, or the acquisition of fiducial markers position by means of a digitizer. For these reasons the accuracy of the tracker-free system has been compared with the one obtainable using commercial localizers taking into account not just the precision of the localization system, but also the whole registration procedure required to align the real and virtual world. For this purpose the fiducials registration error FRE has been evaluated.

In the tracker-free system the marker localization, performed by stereoscopic vision routines, does not require user intervention (automatic segmentation). To determine FRE the helmet has been placed in several positions of the typical workspace and at each position performed localization and registration with preoperative markers positions (given in the CT reference frame) have been performed. Registrations with a FRE > 10 mm (value which guarantee the success of marker segmentation and determined empirically watching debug crosses placed in correspondence of the estimated positions on the grabbed images), have been automatically removed from the FRE evaluation set.

To determine FRE in using commercial localizers, 5 people digitized for 10 times, using different localization systems (FASTRAK Polhemus, NDI Aurora, NDI Optotrak Certus), the 3 fiducial markers consisting in signs on the phantom. Registering markers position acquired by the subjects, with markers positions given in the CT reference frame, the FRE evaluation for each localization systems has been obtained.

The mean values of FRE and standard deviations are reported in Tab. 5.

Localizer	**Mean FRE**	**Std FRE**
Our system	3.47	1.78
FASTRAK Polhemus	3.00	0.41
NDI Aurora	2.72	0.62
NDI Optotrak Certus	2.01	0.54

Tab. 5 Fiducial Registration Error obtained with the binocular system and with commercial localizers.

Regarding the refresh rate, the system is not able to guarantee an enough refresh rate of 25 fps using the entire image for the localization. In order to guarantee a mean 25 fps the zone to

segment for each marker has been reduced to a box of 200x200 pixel, centered on the previous position when FRE was enough (<10 mm). This box dimension allows to find markers in case of typical head movements and the number of frames with wrong localization remain the same that using the entire image: about 2% in an ambulatory room with artificial illumination.

Results demonstrate that using binocular camera localization we introduce an FRE bigger than that one of a commercial localizer. But we have all the advantages of working without a localizer, as described in the introduction. Furthermore, the work presented is only a demonstration of the potentiality of this approach to the problem, and there are some improvements on which we can work in the future, driven by the encouraging results obtained with this simple first system.

Regarding localization precision it is known that stereo cameras localization can reach sub millimetric accuracy depending on the precision of conjugate points determination [82] [83]. In the work just described conjugate points are determined segmenting not uniform felt balls, that guarantee to avoid reflections, and so to have few wrong segmentations, but the determination of their baricenters is not accurate. A useful next step would be the study of materials and shapes for marker fabrication in order to eliminate reflection artifacts and to achieve sub-pixel segmentation accuracy.

5.4 LAPAROSCOPE AUTO LOCALIZATION

This paragraph proposes another solution that allows avoiding the use of an external localization system. In laparoscopic interventions, the position and orientation of the endoscopic camera can be determined with respect to a reference frame fixed to the access ports configuration. Knowing the distances between insertion points, the localization of the endoscopic camera is determined just using information offered by laparoscopic video images.

The proposed solution allows to provide a cheap and tracker-free implementation for a class of computer assisted surgical systems that do not require extremely accurate localization. For example, offering 3D pre-operative model visualization with automatic point of view selection and remote assistance using virtual objects on the laparoscopic monitor. The first assist method concern the viewing of “3D maps”, generated from pre-operative radiological

images (like shown in the chapter 3.0), where the selection of a viewpoint similar to that one of the laparoscope does not require manual work. The viewpoint is adjusted every time that the real camera is moved. This application requires the localization of the camera respect to a reference frame fixed to the patient. Remote assistance can be offered by an expert surgeon moving virtual objects on the laparoscopic monitor. Many remote assistance approaches have been proposed, using videoconference or 2D signs on the images. In this case, 3D virtual objects can be added on the screen, as they would be in the real scene (respecting distances and angles) increasing visual perception and communication respect to the simple 2D overlay. For this application it is necessary to know the relation between real and virtual world.

In both cases a localization accuracy as elevated as for mixed-reality visualization or surgical navigation purposes is not required. Consequently we one can think to localize the camera respect to the scene, like in 5.3, elaborating camera images. Though the insertion of structured artificial markers directly in the abdomen (simple to recognize by machine vision routines) would provide a simple solution to the problem, it is not reasonably acceptable for a real surgical scenario. To overcome this constraint the required information is recovered by artificial objects intrinsically present in laparoscopic images: surgical instruments. They can not be used as a steady reference frame, because they move in the scene, but passing through quite steady points, like access ports in the abdominal wall, the problem can be solved.

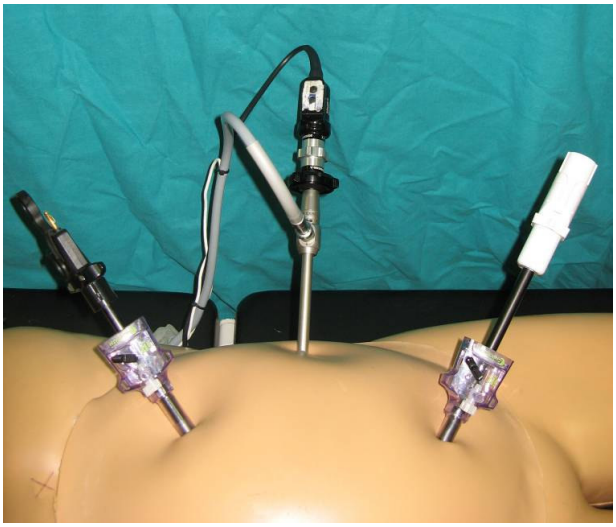

Fig. 37 Typical instruments configuration used in laparoscopy: camera in the middle and two surgical instruments on the sides. The 3 access ports can be used as reference frame.

Any triple of not aligned points, like the access ports in a typical laparoscopic instruments configuration (Fig. 37), allows to define an orthogonal reference frame, which can be used as reference frame for the localization of the camera.

During laparoscopic interventions, camera movements are minor in respect to instruments movements. The cameraman (human or robotic) places the laparoscope in a fixed position to allow the surgeon to view the scene for the current task. The surgeon performs the task using instruments and after that asks the cameramen to move the camera in a new position, for a new surgical task or for exploration purposes. Therefore the laparoscope can be considered steady in a time interval, and instruments positions can be referred to a reference frame fixed on the camera [84] [85]. Insertion point of an instrument could be calculated determining its 3D pose in various images and computing shaft axes intersection [86] using the least square approach. Surgical instruments 3D pose determination, using monoscopic camera, is a difficult task. However it can be calculated adding recognizable markers on instruments [86] or taking into consideration instruments of cylindrical shape, by means of colored strips [87] or elaborating directly instruments projections [88]. The first two solutions require to modify surgical instruments while the third one to work with cylindrical shapes (a situation not always true, for example with opened scissors). In any case three instruments in the scene (for defining a reference frame) would be required, while in nearly every laparoscopic interventions only two instruments are used. The solution proposed determines three insertion points (one camera and two instruments) without the need to compute 3D surgical instruments poses.

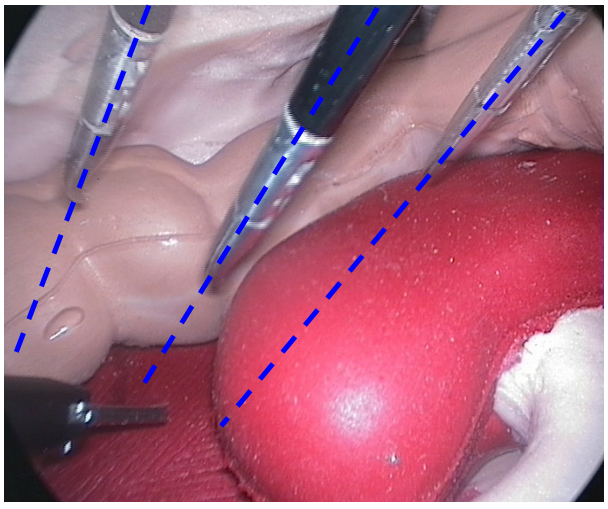

Fig. 38 Image composed by 3 frames of a laparoscopic video with fixed camera and a moving instrument. The projections of instrument axis, represented with blue lines, are constrained to pass through a point representing the projection (on the image plane) of the insertion point (on the abdominal wall).

The projections of instrument axis on the image plane (projection lines), which can be simply determinate using HSV color space and Hough transform [87], are constrained to pass through the projection of the insertion point on the image plane [84] (Fig. 38). The insertion point projection on the image plane can be calculated as the barycentre of the intersection of couples of projection lines, for each instrument. It allows (after camera calibration) to determine the direction of the insertion point in the camera reference frame (Fig. 39).

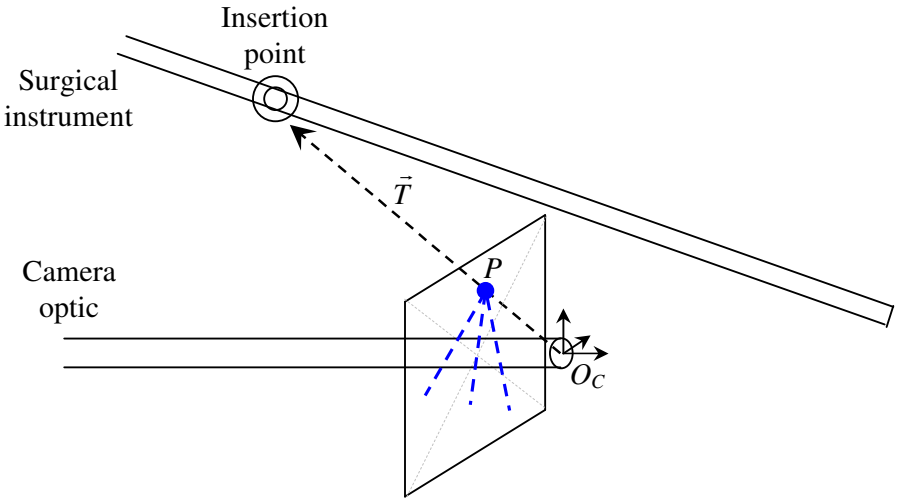

Fig. 39 The projections of instrument axis (blue lines) allow to calculate the projection of a generic insertion point on the image plane *P*, which allows to determinate the direction of the vector $\vec{T}$ representing the insertion point in the camera reference frame fixed on O_C.

Therefore versors $\hat{T}_l$ and $\hat{T}_r$, representing respectively the direction of the left and the right instrument insertion point, can be determinated. The versor $\hat{T}_c$, representing the direction of the camera insertion point, lies on the Z axes of the camera reference frame (using 0 degree optic). The geometrical relation between $\hat{T}_l$, $\hat{T}_r$ and $\hat{T}_c$ can be represented in the following figure:

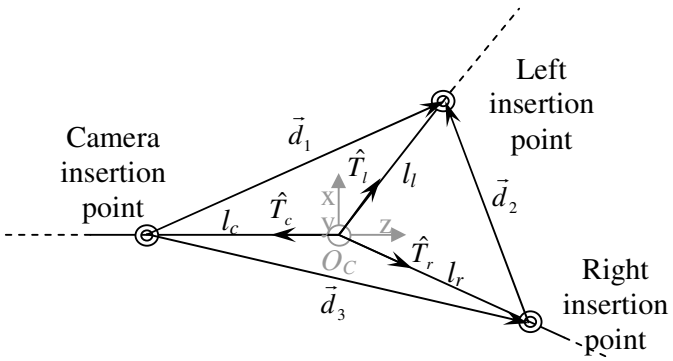

Fig. 40 Geometric relations involved in the insertion points configuration.

and in formula as:

$$\begin{aligned} l_l\hat{T}_l - l_c\hat{T}_c &= \vec{d}_1 \\ l_r\hat{T}_r - l_c\hat{T}_c &= \vec{d}_3 \\ l_l\hat{T}_l - l_r\hat{T}_r &= \vec{d}_2 \end{aligned} \qquad (21)$$

where l_c, l_l and l_r represent the distances of the insertion points from the camera origin, which have to be chosen in order to guaranty the distances between access ports D_1, D_2 and D_3. The tetrahedral configuration allows to determine univocally l_c, l_l and l_r and consequently, having $\hat{T}_c$, $\hat{T}_l$ and $\hat{T}_r$, to localize the access ports respect to the camera (and vice versa).

System evaluation was performed on a trunk phantom for laparoscopy by Sawbones (www.sawbones.com). The localization accuracy depends on the instruments configuration and on their movements. Its mean value has been estimated, in typical conditions, along $\hat{T}_c$, $\hat{T}_l$ and $\hat{T}_r$, in about 1 cm, and evaluated by surgeons enough for the specific application.

PART 3: RESULTS AND CONCLUSIONS

6.0 RESULTS

Patient specific 3D models have been generated and used by physicians for radiological exam visualization and surgical planning in some interventions performed at Cisanello Hospital in Pisa. A liver model generated for a patient with a Klatskin tumor has been used for the surgical planning, the evaluation of the remaining liver functionalities (correlated to the volume removed), and as map in the surgical room during a right epatectomy. Complete navigation functionalities have been clinically tested, tracking surgical instruments by means of a localizer, during a laparoscopic distal pancreatectomy. Experiences in the surgical room shown that the 3D reconstruction of the anatomical structures is a remarkable aid for preoperative surgical planning, providing the surgeon a complete knowledge about the patient anatomy. The surgeon orientation during the intervention is enhanced by virtual views that allow to see the surgical scenario from various viewpoints. The navigation system is also interesting for educational purposes. In fact, students can benefit from additional views giving an interpretation of laparoscopic images and understanding the phases of the intervention in the surgical room. The integration of the navigation system in the OR has not caused discomfort to the surgical staff, and intervention duration has not been affected by the initial setup (about 2 minutes of overload time due to correct repositioning and patient registration).

Regarding registration of the virtual to real anatomy, the various categories of sources of error have been investigated. Experiments for registration error evaluation on external and internal targets have been performed. A rigid body registration to use directly in the surgical room has been designed and tested. The registration accuracy obtained using rigid body approaches allows to develop a whole class of CAS systems for the treatment of the abdominal district where high accuracy is not a critical aspects (such as passive systems providing image guidance), further the developed rigid body registration method can be used as starting point for more accurate registration based on elastic and deformable approaches.

Surgeons evaluated that the mixed-reality is very useful in approaching the target of the intervention providing the same benefits of a GPS for car drivers. The projection model of the virtual camera has been analytically adapted to the real one, and calibration routines, for point of view localization, respect to a sensorized frame, have been developed, obtaining a

perfect alignment between real and virtual view. The alignment accuracy can decrease in surgical scenario, because very little movements on cameras (sensor, optic and sensorized frame) can cause misalignments between real and virtual scene. Images elaboration allowed to develop useful solutions that do not require the use of an external localization system, where camera position is auto-localized in the scene elaborating video images.

7.0 CONCLUSION

Results provided by clinical experimentation and validation of the developed prototypes highlight the relevance of patient specific 3D models and define their role in the enhancement of the social welfare and in the patients' health safeguarding.

The work described in this thesis demonstrates that patient specific 3D models of abdominal organs can be obtained processing medical dataset and used for diagnostic, planning and surgical purposes. Their generation can be performed using the proposed segmentation procedure, coupled with the developed software tool. Future advances, in terms of new organs to segment and in terms of total segmentation time reduction, can be done extending the consideration applied in the current version of the segmentation procedure.

The integration of systems based on 3D models in the surgical room or in the ambulatory for the treatment of disease in daily clinical practice, imposes the consideration of some requirements. Surgery of the future, based on computer aided interventions, requires the integration of localization systems in surgical rooms. This should be done permanently in order to avoid logistics problems and expensive loose of time in devices arrangement and system setup before the beginning of the surgical procedure. On the other hand we can avoid to use an external localization system, in function of clinical requirements, like in the proposed tracker-free stereoscopic see-through and in the laparoscope auto-localization method. Another relevant issue is the integration of sensors, providing information concerning the surgical tools and the patient, and actuators, providing controllable actions, inside a next generation of surgical instruments (based on mechatronics and robotics principles). Even if, the registration error obtained with the current rigid registration method is not enough for the execution of robotic intervention tasks, there are other applications, as the mixed-reality for diagnostic and navigation purposes, where it guaranty enough precision. Further the potential applications can be increased, reducing the source of errors of the registration, as described, and acquiring intra-operative information, elaborating camera images, as shown.

The actuation of the clinical scenario of the future (based on information and robotic technologies) requires the integration in the radiological departments and in the operative rooms of CAD/CAS professional figures having high technical background (engineering and

IT) and high competences in medical and surgical procedures. The definition of their role in clinical workflows can enhance drastically the fusion and exploitation of preoperative, intraoperative and postoperative medical information, for diagnosis, communication (between radiological and therapeutic departments), planning and surgical execution, training and simulation purposes, and consequently can allow the development of new useful therapeutic strategies in line with the running technological development.

APPENDIX A: SVD REGISTRATION

If we have two point clouds $p_i=\{x_i, y_i, z_i\}^T$ and $p'_i=\{x_i, y_i, z_i\}^T$, $i=1,2,\ldots n$ to register, we can consider p_i and p'_i as 3x1 column matrixes, which can be related by:

$$p'_i = Rp_i + T + N_i$$

where R is a 3x3 rotation matrix , T a translation vector and N noise. Least square solution means to minimize:

$$\Sigma^2 = \sum_{i=1}^{N} \left\| p'_i - (Rp_i + T) \right\|^2$$

As demonstrated in [60] this problem can be decoupled because, if the least square solution to *(22)* is $\hat{R}$ and $\hat{T}$, the barycentre of $\{p_i\}$ and $\{\hat{R}p'_i + \hat{T}\}$ coincide. A demonstration can be the following. Let:

$$p''_i = \hat{R}p_i + \hat{T}$$

we can rewrite the optimization problem as: find p''_i that minimize

$$\Sigma^2 = \sum_{i=1}^{N} \left\| p'_i - p''_i \right\|^2$$

subject to the rigid constrain:

$$\left\| p''_i - p''_{ji} \right\|^2 = \left\| p_i - p_{ji} \right\|^2 \forall i, j$$

Using the lagrangian multipliers we can write this new problem in the compact form:

$$F = \sum_{i=1}^{N} \left\| p'_i - p''_i \right\|^2 + \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_{ij} \left(\left\| p''_i - p''_j \right\|^2 - \left\| p_i - p_j \right\|^2 \right)$$

In order to minimize F we can differentiate respect to x''_i obtaining:

$$\frac{dF}{dx''_i} = 0 = 2(x''_i - x'_i) + \sum_{j=1}^{n} \lambda_{ij} 2\left(x''_i - x''_j\right)$$

Summering over all i we obtain:

$$\sum_{i=1}^{N} \left\| x''_i - x'_i \right\| = 0 \Rightarrow \sum_{i=1}^{N} \left\| x''_i \right\| = \sum_{i=1}^{N} \left\| x'_i \right\|$$

Similarly for the y and z components, so if we calculate the barycentres of p_i, p''_i and p'_i :

$$p = \frac{1}{N} \sum_{i=1}^{N} p_i \ , \quad p' = \frac{1}{N} \sum_{i=1}^{N} p'_i \ , \quad p'' = \frac{1}{N} \sum_{i=1}^{N} p''_i$$

we obtain:

$$p' = p'' = \hat{R}p + \hat{T}$$

So, the matrix $\hat{R}$ can be found solving the following problem:

$$q_i = p_i - p$$
$$q'_i = p'_i - p'$$
$$\min\left(\Sigma^2 = \sum_{i=1}^{N} \left\| q'_i - Rq_i \right\|^2\right)$$

Espanding Σ^2 , how in [54] we obtain:

$$\begin{aligned}\Sigma^2 &= \sum_{i=1}^{N} (q_i' - Rq_i)^t (q_i' - Rq_i) \\ &= \sum_{i=1}^{N} (q_i'^t q_i' + q_i^t R^t R q_i - q_i'^t R q_i - q_i^t R^t q_i') \\ &= \sum_{i=1}^{N} (q_i'^t q_i' + q_i^t q_i - 2 q_i'^t R q_i)\end{aligned}$$

Therefore, minimizing Σ^2 is equivalent to maximize:

$$\begin{aligned}F &= \sum_{i=1}^{N} (q_i'^t R q_i) \\ &= Trace\left(\sum_{i=1}^{N} (R q_i q_i'^t)\right) \\ &= Trace(RH)\end{aligned}$$
$$where \qquad H = \sum_{i=1}^{N} (q_i q_i'^t)$$

H can be decomposed using the singular value decomposition (SVD) as:

$$H = USV^t$$

Where U and V are 3X3 orthonormal matrixes and S is a 3X3 diagonal matrix containing singular values.

Choosing $R=VU^t$ we maximize *F*, because:

$$Trace(RH) = Trace(VU^t USV^t) = Trace(VSV^t)$$

and if we apply any other rotation *B* to *R*:

$$Trace(BRH) = Trace(BVSV^t) = Trace((BV)SV^t)^t = Trace(V^t B^t VS)$$

Now, using the column vectors v_i of *V* and the diagonal value s_i of *S*:

$$Trace(BRH) = \sum_i (v_i^t B v_i s_i) = \sum_i (s_i v_i^t B v_i)$$

for each i, following the Schwarz inequality:

$$s_i v_i^t (B v_i) \le s_i \sqrt{(v_i^t v_i)(v_i^t B^t B v_i)} = s_i v_i^t v_i$$

where $s_i v^t_i v_i$ is each diagonal value of VSV^t , and so, any other additional rotation B can not increase the value F (to maximize).

Note that choosing $\hat{R} = VU^t$ we obtain an orthogonal matrix, but can not be orthonormal, having a determinant =-1. This situation depends on the noise and points coplanar configurations, but however we can invert the sign of the third column of V obtaining a rotation.

Finally $\hat{T}$ can be set, following the demonstration on problem decoupling, as:

$$\hat{T} = p' - \hat{R} p$$

APPENDIX B: PCA REGISTRATION

After the alignment of the barycentres of the points clouds $p_i=\{x_i, y_i, z_i\}^T$ and $p'_i=\{x_i, y_i, z_i\}^T$, as shown in the previous appendix:

$$q_i = p_i - p$$
$$q'_i = p'_i - p'$$

we can align the inertial axis of the new iso-barycentre points clouds composed with all q_i and q'_i. Considering each q_i and q'_i as a unitary material point we can calculate the corresponding inertia matrix I and I'.

From wikipedia:

> "For a rigid object of N point masses m_k, the inertia matrix is given by:
>
> $$\mathbf{I} = \begin{bmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{bmatrix}.$$
>
> Its components are defined as
>
> $$I_{ij} \stackrel{\text{def}}{=} \sum_{k=1}^{N} m_k (r_k^2 \delta_{ij} - r_{ki} r_{kj})$$
>
> Where:
>
> i, j equal 1, 2, or 3 for x, y, and z, respectively,
>
> r_k is the vector to the mass k from the point about which the tensor is calculated, and δ_{ij} is the Kronecker delta."

I can be expressed using the covariance matrix C as follow:

$$Q = [q_1 \quad .. \quad q_n]$$
$$Q' = [q_1' \quad .. \quad q_n']$$
$$C = Q \cdot Q^T$$
$$C' = Q' \cdot Q'^T$$
$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} trace(C) - C$$
$$I' = ...$$

perhaps if C is diagonal the inertial matrix I will be diagonal (principal inertial axes aligned with the reference frame), and if C' is diagonal I' will be diagonal.

C and C' can be decomposed using principal component analysis (PCA) decomposition:

$$C = U \cdot S \cdot U^T$$
$$C' = U' \cdot S' \cdot U'^T$$

where S and S' are diagonal and U and U' are orthogonal matrixes.

Working with the definition of C we obtain:

$$C = Q \cdot Q^T$$
$$C = U \cdot S \cdot U^T$$
$$U^{\mathrm{T}} \cdot C \cdot U = S$$
$$(U^{\mathrm{T}} \cdot Q) \cdot (Q^T \cdot U) = (U^{\mathrm{T}} \cdot Q) \cdot (U^{\mathrm{T}} \cdot Q)^T = S$$

So U^T is the rotation that align the axes of the points cloud contained in Q with the axes of the reference frame. Following the same considerations U'^T is the rotation that align the axes of the points cloud contained in Q' with the axes of the reference frame.

Definitely the principal inertial axes of the points cloud contained in Q can be aligned with the principal inertial axes of the points cloud contained in Q' using the rotation matrix R as following defined:

$$R = \left(U'^T\right)^{-1} \cdot U^T = U' U^T$$

We can demonstrate that in case of no noise on data PCA and SVD registrations methods offer (obviously) the same result, allow to estimate the exact rotation matrix R.

Using the PCA method we have to calculate:

$$QQ^T = U \cdot S \cdot U^T$$
$$Q'Q'^T = U' \cdot S' \cdot U'^T$$
$$\hat{R} = U'U^T$$

Considering no noise:

$$QQ^T = U \cdot S \cdot U^T$$
$$Q'Q'^T = RQ(RQ)^T = RQQ^TR^T = (RU)S(U^TR^T)$$
$$\hat{R} = U'U^T = RUU^T = R$$

Using SVD method we have to calculate:

$$QQ'^T = USV^t$$
$$\hat{R} = VU^T$$

Considering no noise and remembering the PCA decomposition of QQ^T:

$$QQ'^T = Q(RQ)^T = QQ^TR^T = USU^TR^T = US(U^TR^T) = US(RU)^T$$
$$\hat{R} = RUU^T = R$$

BIBLIOGRAPHY

[1] L. T. Kohn, J. M. Corrigan, and M. S. Donaldson, Eds., *To Err is Human*. IoM, 2000.

[2] H. Ouellette, A. Kassarjian, and T. McLoud, "Teaching the art of verbal consultation," *Journal of the American College of Radiology*, vol. 3(1), pp. 9–10, 2006.

[3] G. Megali, ""more than" a navigation system for assistance in surgery," Ph.D. dissertation, MiTech Laboratory, Scuola Superiore Sant'Anna, Pisa, Italy, January 2002.

[4] P. Dario, B. Hannaford, and A. Menciassi, "Smart surgical tools and augmenting devices," *IEEE Transactions on Robotics and Automation*, vol. 19, no. 5, pp. 782–792, October 2003.

[5] K. K. Herfarth, J. Debus, F. Lohr, M. L. Bahner, P. Fritz, A. Höss, W. Schlegel, and M. F. Wannenmacher, "Extracranial stereotactic radiation therapy: set-up accuracy of patients treated for liver metastases," *International journal of radiation oncology, biology, physics*, vol. 46(2), pp. 329–35, 2000.

[6] D. Stoianovici, D. Song, D. Petrisor, D. Ursu, D. Mazilu, M. Muntener, M. Schar, and A. Patriciu, ""mri stealth" robot for prostate interventions," *Minim Invasive Ther Allied Technol*, vol. 16(4), pp. 241–8, 2007.

[7] G. Fichtinger, J. P. Fiene, C. W. Kennedy, G. Kronreif, I. Iordachita, D. Y. Song, E. C. Burdette, and P. Kazanzides, "Robotic assistance for ultrasound-guided prostate brachytherapy," in *Medical Image Analysis*, vol. 12, no. 5, 2008, pp. 535 – 545.

[8] P. Milgram and F. Kishino, "A taxonomy of mixed reality visual displays," *IEICE transactions on information and systems*, vol. 77, no. 12, pp. 1321–1329, 1994.

[9] J. H. Kaspersen, E. Sjolie, J. Wesche, Asland, J. Lundbom, A. Odegard, F. Lindseth, and T. A. Nagelhus-Hernes, "Three-dimensional ultrasound-based navigation combined with preoperative ct during abdominal interventions: A feasibility study," *CardioVascular and Interventional Radiology*, vol. V26, no. 4, pp. 347–356, 2003.

[10] F. Cavallo, "Surgical gesture analysis:systems and methods for biomechanical modelling of surgeon in minimally invasive surgey," Ph.D. dissertation, Pisa, Italy, 2007.

[11] S. Sinigaglia, "Surgical gesture analysis:a machine learning approach.from acquisition to understanding surgical performance," Ph.D. dissertation, Pisa, Italy, 2007.

[12] G. Megali, V. Ferrari, C. Freschi, B. Morabito, F. Cavallo, G. Turini, E. Troia, C. Cappelli, A. Pietrabissa, O. Tonet, A. Cuschieri, P. Dario, and F. Mosca, "Endocas navigator platform: a common platform for computer and robotic assistance in minimally invasive surgery," *The International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 4, pp. 242 – 251, 2008.

[13] W. E. Lorensen and H. E. Cline, "Marching cubes: A high resolution 3D surface construction algorithm," in *SIGGRAPH '87: Proceedings of the 14th annual conference on Computer graphics and interactive techniques*. New York, NY, USA: ACM, 1987, pp. 163–169.

[14] R. A. ROBB, "Biomedical imaging: Past, present and predictions," in *Proceedings of the First International Symposium on Intelligent Assistance in Diagnosis of Multi-Dimensional Medical Image*, 2005.

[15] C. Kirbas and F. Quek, "A review of vessel extraction techniques and algorithms," in *ACM Comput. Surv.* New York, NY, USA: ACM Press, 2004, vol. 36, no. 2, pp. 81–121.

[16] D. L. Pham, C. Xu, and J. L. Prince, "A survey of current methods in medical image segmentation," *Annual Review of Biomedical Engineering*, vol. 2, pp. 315–337, 2000.

[17] A. Bornik, R. Beichel, and D. Schmalstieg, "Interactive editing of segmented volumetric datasets in a hybrid 2D/3D virtual environment." *Proceedings of the ACM symposium on Virtual reality software and technology*, pp. 197–206, 2006.

[18] P. M. A. van Ooijen, R. Wolf, A. Schenk, D. B. Rouw, M. Slooff, and M. O. H.-O. Peitgen, "Recent developments in organ-selective reconstruction and analysis of multiphase liver ct," *Imaging Decisions MRI*, vol. 7, pp. 37–43, 2003.

[19] D. V. Amin, T. Kanade, B. Jaramaz, A. M. DiGioia, C. Nikou, R. S. LaBarca, and J. E. MoodyJr., "Calibration method for determining the physical location of the ultrasound image plane." in *MICCAI*, ser. Lecture Notes in Computer Science, W. J. Niessen and M. A. Viergever, Eds., vol. 2208. Springer, 2001, pp. 940–947.

[20] J. B. A. Maintz and M. A. Viergever, "A survey of medical image registration," *Medical Image Analysis*, vol. 2, no. 1, p. 1–37, 1998.

[21] R. Adams and L. Bischof, "Seeded region growing," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 16, no. 6, pp. 641–647, Jun 1994.

[22] J. Fan, G. Zeng, M. Body, and M.-S. Hacid, "Seeded region growing: an extensive and comparative study," *Pattern Recognition Letters*, vol. 26, no. 8, pp. 1139–1156, June 2005.

[23] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International Journal of Computer Vision*, vol. 1(4), pp. 321–331, 1998.

[24] J. Suri, K. Liu, S. Singh, S. Laxminarayana, and L. Reden, "Shape recovery algorithms using level sets in 2-d/3-d medical imagery: A state-of-the-art review," *IEEE Trans. in Information Technology in Biomedicine*, vol. 6, No. 1, pp. 8–28, 2002.

[25] T. Rohlfing, R. Brandt, R. Menzel, D. B. Russakoff, and C. R. Maurer, Jr., "Quo vadis, atlas-based segmentation?" in *The Handbook of Medical Image Analysis – Volume III: Registration Models*, J. Suri, D. L. Wilson, and S. Laxminarayan, Eds. New York, NY: Kluwer Academic / Plenum Publishers, Aug. 2005, ch. 11, pp. 435–486.

[26] J. Costa, H. Delingette, S. Novellas, and N. Ayache, "Automatic segmentation of bladder and prostate using coupled 3D deformable models." in *MICCAI*, 2007.

[27] R. Shimizu, A.and Ohno, T. Ikegami, H. Kobatake, S. Nawano, and D. Smutek, "Multi-organ segmentation in three dimensional abdominal ct images," *IJCARS*, vol. 1; SUPP/1, pp. 76–77, 2006.

[28] A. Shimizu, T. Kimoto, D. Furukawa, H. Kobatake, and K. Shinozaki, "Pancreas segmentation in three phase abdominal ct volume data," in *Proceedings of CARS*, 2008.

[29] D. Selle, B. Preim, A. Schenk, and H. Peitgen, "Analysis of vasculature for liver surgical planning," *IEEE Transactions on Medical Imaging*, vol. 21 (11), pp. 1344–1357, 2002.

[30] L. Soler, N. Ayache, S. Nicolau, X. Pennec, C. Forest, H. Delingette, D. Mutter, and J. Marescaux, "Virtual reality, augmented reality and robotics in surgical procedures of the liver," in *Perspectives in Image-guided Surgery. Proceedings of the Scientific Workshop on Medical Robotics, Navigation and Visualization (MRNV) 2004*, T. M. Buzug and T. C. Lueth, Eds. RheinAhrCampus Remagen, Germany: World Scientific, March 11-12 2004, pp. 476–484.

[31] J. Costa, "Segmentation of anatomical structures of the lower abdomen using 3D deformable models." PhD Thesis, École Nationale Supérieure des Mines de Paris, March 2008.

[32] J. L. Foo, "A survey of user interaction and automation in medical image segmentation methods, technical report isu-hci-2006-2," Iowa State Universtity - Human Computer Interaction, Tech. Rep., 2006.

[33] L. Ibáñez, W. Schroeder, L. Ng, J. Cates, and R. Hamming, "The itk software guide," 2005.

[34] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, pp. 629–639, 1990.

[35] Carlson, Felmlee, Bender, Ehman, Classic, Hu, and Hoskin, "Intermittent-mode ct fluoroscopy-guided biopsy of the lung or upper abdomen with breath-hold monitoring and feedback system development and feasibility," *Radiology*, vol. 229, pp. 906–912, 2003.

[36] B. Zitova and J. Flusser, "Image registration methods: a survey," *Image and Vision Computing*, vol. 21, no. 11, pp. 977–1000, October 2003.

[37] L. Ramirez, N. Durdle, and V. Raso, "Medical image registration in computational intelligence framework: a review," in *Electrical and Computer Engineering, 2003. IEEE CCECE 2003. Canadian Conference on*, vol. 2, 4-7 May 2003, pp. 1021–1024.

[38] M. I. Miga, K. D. Paulsen, J. M. Lemery, S. D. Eisner, A. Hartov, F. E. Kennedy, and D. W. Roberts, "Model-updated image guidance: Initial clinical experiences with gravity-induced brain deformation." *IEEE Trans. Med. Imaging*, vol. 18, no. 10, pp. 866–874, 1999.

[39] O. Commowick, V. Arsigny, A. Isambert, J. Costa, F. Dhermain, F. Bidault, P.-Y. Bondiau, N. Ayache, and G. Malandain, "An efficient locally affine framework for the smooth registration of anatomical structures," *Medical Image Analysis*, vol. 12, no. 4, pp. 427–441, 2008.

[40] M. Feuerstein, T. Mussack, S. M. Heining, and N. Navab, "Registration-free laparoscope superimposition for intra-operative planning of liver resection," in *3rd Russian-Bavarian Conference on Biomedical Engineering*, Erlangen, Germany, July 2007, pp. 88–92.

[41] D. Hawkes, D. Barratt, J. Blackall, A. Chandler, J. McClelland, and G. Penney, "Computational models in image guided interventions," in *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*, 2005, pp. 7246–7249.

[42] B. Olbricha, J. Traubb, S. Wiesnerb, A. Wicherta, H. Feussnera, and N.Navabb, "Respiratory motion analysis: Towards gated augmentation of the liver," in *CARS 2005 Computer Assisted Radiology and Surgery, 19st International Congress and Exhibition*, 2005.

[43] J. M. Blackall, S. Ahmad, M. E. Miquel, J. R. McClelland, D. B. Landau, and D. J. Hawkes, "Mri-based measurements of respiratory motion variability and assessment of imaging strategies for radiotherapy planning." *Phys Med Biol*, vol. 51, no. 17, pp. 4147–4169, Sep 2006.

[44] J. R. McClelland, J. M. Blackall, S. Tarte, A. C. Chandler, S. Hughes, S. Ahmad, D. B. Landau, and D. J. Hawkes, "A continuous 4d motion model from multiple respiratory cycles for use in lung radiotherapy." *Med Phys*, vol. 33, no. 9, pp. 3348–3358, Sep 2006.

[45] H. Delingette, X. Pennec, L. Soler, J. Marescaux, and N. Ayache, "Computational models for image-guided robot-assisted and simulated medical interventions," *Proceedings of the IEEE*, vol. 94, no. 9, pp. 1678–1688, Sept. 2006.

[46] M. Riboldi, G. Baroni, M. Spadea, B. Tagaste, C. Garibaldi, P. Fossati, R. Orecchia, and A. Pedotti, "Implementation of fem simulation techniques for organ motion modeling in prostate cancer radiotherapy," in *Radiotheraphy Oncology*, 2006.

[47] N. Ayache, J.-P. Boissel, S. Brunak, G. Clapworthy, G. Lonsdale, J. Fingberg, A. Frangi, G. Deco, P. Hunter, P. Nielsen, M. Halstead, R. Hose, I. Magnin, F. Martin-Sanchez, P. Sloot, J. Kaandorp, A. Hoekstra, S. V. S. Jan, and M. Viceconti, "Towards virtual physiological human: Multilevel modelling and simulation of the human anatomy and physiology," Virtual Physiological Human: White paper, EC - DG INFSO and DG JRC, 2006.

[48] M. Scheuering, C. Rezk-Salama, H. Barfufl, A. Schneider, and G. Greiner, "Augmented reality based on fast deformable 2D-3D registration for image-guided surgery," in *Proc. SPIE Vol. 4681, p. 436-445, Medical Imaging 2002: Visualization, Image-Guided Procedures, and Display, Seong K. Mun; Ed.*, ser. Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference, S. K. Mun, Ed., vol. 4681, May 2002, pp. 436–445.

[49] G. Baroni, C. Garibaldi, M. Riboldi, M. Spadea, G. Catalano, B. Tagaste, G.Tosi, R. Orecchia, and A. Pedotti, "3D optoelectronic analysis of interfractional patient setup variability in frameless extracranial stereotactic radiotherapy," *Int J Radiat Oncol Biol Phys*, vol. 64(2), pp. 635–642, 2006.

[50] M. van Herk, "Errors and margins in radiotherapy," *Seminars in Radiation Oncology*, vol. 14, p. 52–64, 2004.

[51] J. Ayoub, R. Cohendy, M. Dauzat, R. Targhetta, J. E. D. la Coussaye, J. M. Bourgeois, M. Ramonatxo, C. Prefaut, and L. Pourcelot, "Non-invasive quantification of diaphragm kinetics using m-mode sonography." *Can J Anaesth*, vol. 44, no. 7, pp. 739–744, Jul 1997.

[52] C. Song, A. Alijani, T. Frank, G. Hanna, and A. Cuschieri, "Mechanical properties of the human abdominal wall measured in vivo during insufflation for laparoscopic surgery." *Surg Endosc*, vol. 20(6), pp. 987–990, 2006.

[53] W. Segars, D. Lalush, and B. Tsui, "Modeling respiratory mechanics in the mcat and spline-based mcat phantoms," *Nuclear Science, IEEE Transactions on*, vol. 48, p. 89–97, 2001.

[54] K. S. Arun, T. S. Huang, and S. D. Blostein, "Least-squares fitting of two 3-d point sets," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 9, no. 5, pp. 698–700, 1987.

[55] P. J. Besl and N. D. McKay, "A method for registration of 3-d shapes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 2, pp. 239–256, 1992.

[56] H. ad Kearns, Ellis, Sprinkle, Cullen, Smith, and Stieber, "Reducing set-up uncertainty in the elekta stereotactic body frame using stealthstation software," *Technology in cancer research & treatment*, vol. 6, pp. 181–6, 2007.

[57] D. J. Hawkes, G. P. Penney, D. Atkinson, D. C. Barratt, J. M. Blackall, T. J. Carter, W. R. Crum, J. McClelland, C. Tanner, S. Tarte, and M. White, "Motion and biomechanical models for image-guided interventions," in *ISBI*, 2007, pp. 992–995.

[58] B. Brendela, S. Winterb, A. Rickc, M. Stockheimd, and H. Ermert, "Bone registration with 3d ct and ultrasound data sets," *International Congress Series*, vol. 1256, pp. 426–432, 2003.

[59] J. M. Fitzpatrick, J. B. West, and C. R. J. Maurer, "Predicting error in rigid-body point-based registration," *IEEE Transactions on Medical Imaging*, vol. 17, no. 5, pp. 694–702, Oct 1998.

[60] T. S. Huang, S. D. Blostein, and E. A. Margerum, "Least-squares estimation of motion parameters from 3-d point correspondences," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 1986.

[61] G. Calafiore, "On 3-d point set matching with mae and sae cost criteria," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 38(2), pp. 443–450, 2007.

[62] V. Ferrari, G. Megali, C. Cappelli, E. Troia, F. Cavallo, and A. Pietrabissa, "Improving daily clinical practice with 3D patient-specific anatomical models: limits, methodologies and our experience," in *6th International Conference on the Management of Healthcare & Medical Technology - HCTM*. Scuola Superiore Sant'Anna, Pisa, Italy, 3-5 October 2007 2007.

[63] P. Cosi, V. Ferrari, E. Magno, G. Perin, G. Tisato, and C. Zmarich, "Greta e lucia: due realistiche facce parlanti animate mediante un nuovo modello di coarticolazione," pp. 127–134, 2003.

[64] F. I. Parke and K. Waters, *Computer facial animation*. Natick, MA, USA: A. K. Peters, Ltd., 1996.

[65] J. H. Shuhaiber, "Augmented reality in surgery," *Archive of surgery*, vol. 139, pp. 170–174, 2004.

[66] C. Bichlmeier, T. Sielhorst, S. M. Heining, and N. Navab, "Improving depth perception in medical ar. a virtual vision panel to the inside of the patient," in *Bildverarbeitung für die Medizin*, 2007.

[67] L. Johnson, E. Philip, G. Lewis, and D. Hawkes, "Depth perception of stereo overlays in image-guided surgery," in *Medical Imaging, Proceedings of the SPIE, Volume 5372, pp. 263-272.*, 2004.

[68] H. Hua, P. Krishnaswamy, and J. P. Rolland, "Video-based eyetracking methods and algorithms in head-mounted displays," *Optics Express*, vol. 14, pp. 4328–4350.

[69] E. C. Lee and K. R. Park, "A robust eye gaze tracking method based on a virtual eyeball model," *Machine Vision and Applications*, 2008.

[70] Y. Genc, F. Sauer, F. Wenzel, M. Tuceryan, and N. Navab, "Optical see-through hmd calibration: A stereo method validated with a video see-through system," in *International Symposium for Augmented Reality*, 2000.

[71] Y. Baillot and S. J. Julier, "A tracker alignment framework for augmented reality," in *In Proc. Second IEEE and ACM International Symposium on Mixed and Augmented Reality*. IEEE, 2003, pp. 142–150.

[72] *Halcon Application Guide 7.1*. MVTex Software GmbH, 2005.

[73] C. Gao and N. Ahuja, "Single camera stereo using planar parallel plate," in *Pattern Recognition, 17th International Conference on (ICPR'04) Volume 4*, 2004, pp. 108–111.

[74] Baumhauer, Simpfendörfer, Schwarz, Seitel, Müller-Stich, Gutt, R. Meinzer, and I. Wolf, "Soft tissue navigation for laparoscopic prostatectomy: evaluation of camera pose estimation for enhanced visualization," in *Medical Imaging 2007: Visualization and Image-Guided Procedures. Proceedings of the SPIE, Volume 6509, pp. 650911*, 2007.

[75] C. Wengert, L. Bossard, A. Haeberling, C. Baur, G. Székely, and P. C. Cattin, "Endoscopic navigation for minimally invasive suturing," ser. Lecture Notes in Computer Science, N. Ayache, S. Ourselin, and A. J. Maeder, Eds., vol. 4792. Springer, 2007, pp. 620–627.

[76] C. Doignon, F. Nageotte, B. Maurin, and A. Krupa, "Model-based 3-d pose estimation and feature tracking for robot assisted surgery with medical imaging," in *From Features to Actions - Unifying Perspectives in Computational and Robot Vision, Workshop at the IEEE Int. Conf. on Robotics and Automation*, D. Kragic, Ed., Roma, Italy, April 2007.

[77] R. Hartley and A. Zisserman, "Multiple view geometry in computer vision," 2004.

[78] S. Speidel, M. Kleinert, G. Sudra, C. Müller-Stich, B. P.and Gutt, and R. Dillmann, "3d-reconstruction from endoscopic image sequences for intraoperative assistance," in *CARS*, 2008.

[79] C. McGlone, E. Mikhai, and J. Bethel, *Manual of Photogrammetry, fifth ed.* American Society of Photogrammetry and Remote Sensing, 2004.

[80] G. Wei, K. Arbter, and G. Hirzinger, "Real time visual servoing for laparoscopic surgery." *IEEE Engineering in Medicine and Biology*, vol. 16, pp. 40 – 45, 1997, lIDO-Berichtsjahr=1997,;. [Online]. Available: http://elib.dlr.de/28582

[81] V. Kravtchenko, "Thesis title: Tracking color objects in real time," Master's thesis, The University of British Columbia, 1999.

[82] J. Amat, M. Frigola, and A. Casals, "Selection of the best stereo pair in a multi-camera configuration," in *ICRA*, 2002, pp. 3342–3346.

[83] H. Sahabi and A. Basu, "Analysis of error in depth perception with vergence and spatially varying sensing," *Comput. Vis. Image Underst.*, vol. 63, no. 3, pp. 447–461, 1996.

[84] C. Doignon, F. Nageotte, and M. de Mathelin, "The role of insertion points in the detection and positioning of instruments in laparoscopy for robotic tasks," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI*, 2006.

[85] S. Voros, J.-A. Long, and P. Cinquin, "Automatic localization of laparoscopic instruments for the visual servoing of an endoscopic camera holder," in *MICCAI (1)*, 2006, pp. 535–542.

[86] F. Naogette, P. Zanne, C. Doignon, and M. De Mathelin, "Visual servoing-based endoscopic path following for robot-assisted laparoscopic surgery," in *International Conference on Intelligent Robots and Systems (IROS)*. Beijing, China: IEEE, october 2006.

[87] O. Tonet, T. Ramesh, G. Megali, and P. Dario, "Tracking endoscopic instruments without localizer: a shape analysis-based approach," *Computer Aided Surgery*, vol. 12, no. 1, pp. 35–42, january 2007.

[88] C. Doignon and M. De Mathelin, "A degenerate conic-based method for a direct fitting and 3-d pose of cylinders with a single perspective view," in *IEEE Int'l Conf. On Robotics and Automation*, 10-14 April 2007.